UNIVERSITY OF TORONTO

# Assignment 12: UTM Timetable Planner Redesign

Manahil Masroor, Wei Hsuan (Jacob) Lin, Miguel Duterte, Shobhit Srivastava

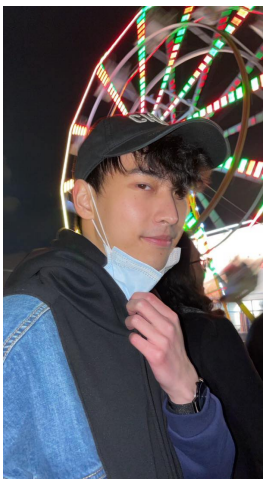
1004351177, 1005094906, 1005327852, 1004947310

CCT480H5 PRA9101

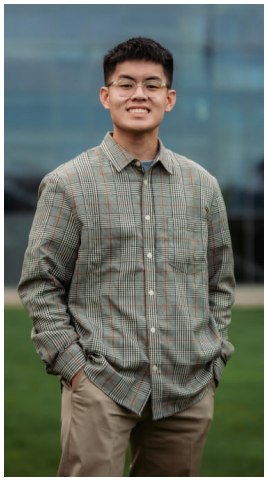
Professor Amna Liaqat, TA Eleen Gong

05/14/2022

Manahil Masroor



Wei Hsuan (Jacob) Lin



Miguel Duterte

Shobhit Srivastava

* All changes were made in accordance with feedback from TA.Changes from our previously submitted assignments are in accordance with the following format:

Changes from previous assignments.

**Justification for changes is **Changed from A11 version 1:** originally had an alt hypothesis focused on accuracy and combined number of clicks.*

## Table of Contents

## Executive Summary

The focus of our project was redesigning the **UTM Timetable planner website.** The main usability fail we noticed with the site was that there is an excessive amount of scrolling when going between adding/adjusting courses and opening up students' planned schedule.

We iterated through different design ideas through the **10 + 10 design ideation** method and came up with the idea of having the website be a split view of the courses and the schedule, as well as courses being separate hexagons in a honey comb sort of design to address the limited space issue and decrease the need of excessive scrolling. Using this idea we created our first iteration for our **low-fi prototype** with annotations and a **user map** to get context later about the flow and which screens to create in our interactive prototype.

With this **interactive Figma prototype**, we conducted **guerilla testing** with 5 users in a "think-aloud" style from users. Using user feedback updates were made to the prototypes such as making it intuitive that hexagons should be clicked and differentiating winter and fall courses by colour. The timetable iteration was then tested in **benchmarking testing** where we benchmarked measurement for tasks such as **searching** for a CCT 4th Year course, **opening lecture times** for a certain course, **adding a course** to the plan and **switching to the winter schedule.** We then used feedback and the results such as number of clicks to open lecture times to update our prototypes again such as adding stroke to hexagons for more accessibility. The data collected from benchmark testing was then contextualized as we identified the measure of centrality and spread.

An **alternative card design** was created based on user feedback which stated that users found the order of courses confusing, so having the courses in cards made a more traditional layout when reading the content as well as address our pain point of excessive scrolling. This provided a contrast to our initial design with a "creative" approach, while our alternative design was more of a "traditional" approach. Our team then produced and **A/B testing protocol** and conducted two rounds of **A/B testing with Design A (the card design) and Design B (our original hex design)** by asking users to do tasks of **locating a specific course (task A)**, **add the course** to the plan **(task B)**, and **remove the course** from the plan **(task C)** for both designs in counterbalanced order. We measured **time taken** and **system usability (SUS)** for task A, **# of clicks** for task B and C, and **overall satisfaction** for all tasks. Our initial phase of A/B testing provided us with solidifying our scenarios, protocols, and minor issues with our prototypes (such as typos) for our final round of testing.

We created our research questions of which design is better in terms of comprehension and satisfaction, our **null hypothesis** stated: 1) There is **no difference** between Design A (alternative card) and Design B (hexagonal Design) in terms of **comprehension** and 2)There is **no difference** between Design A (alternative card) and Design B (hexagonal Design) in terms of **satisfaction rate for each task**, our **alternative hypothesis,** identified **independent**, **dependent**, and **confounding** variables, potential **Type 1** and **Type 2 errors**. We then performed **summary statistics** and **statistical tests** on our A/B testing results with **interpretations** on if our results rejected or failed to reject our null hypotheses. Our team also stated any **limitations** we ran into our testing, design and project overall.

## Assignment 1: Project Idea

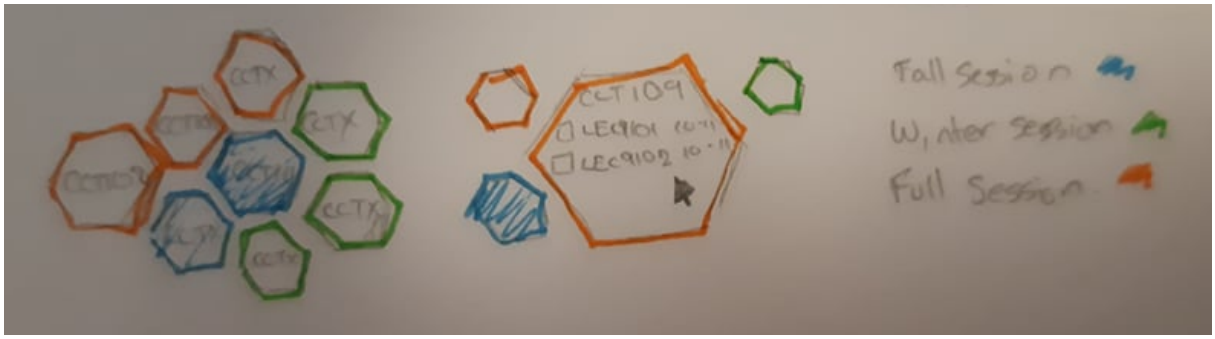
**Problem:**

- **Who:** The problem and the solution to the problem directly affects current lower year students (specifically) at University of Toronto Mississauga (UTM).
- **What:** The stressful process of planning and adjusting courses manually via UTM Timetable Planner.
- **Where:** On the UTM timetable planner website during release of semester class schedules.
- **Why:** Currently students have to manually select each course and section/time of lecture. Conflicts are not clearly indicated on the timetable. Advanced options of providing lectures on certain days are also not provided. Refreshing the page loses all your planning information and students have to start all over if they didn't record the information. Many times students (especially lower year students) have trouble planning and find conflicts in courses through the lack of planning skills students possess or lack of planning features on the website. This often stresses out students.
- **How**: UTM Timetable planner allows input of semester (fall or winter), year of study, and program area (ccit, cs, polisci etc.), advanced options also allow course code, instructor name. Conflicts are just symbolized by coloring the conflict block red. Refreshing the page loses all the information as no sign in or any way of recording is used on the website.

## Assignment 2: Usability Fails & 10 + 10 Designs

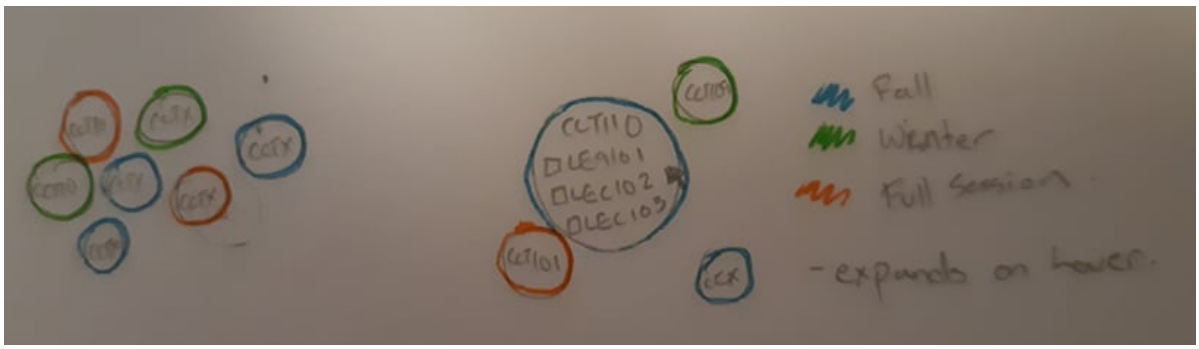
**Pain Point:** The pain point the team determined to be the biggest pain point for users was the excessive amount of scrolling users have to do between adding/adjusting courses and seeing their planned schedule.
**Added a more specific pain point, as our pain points from A2 version 1 focused on too many issues stated in A1. We narrowed it down to only focus and tackle one issue.*

**Part 1 of 10 + 10: Different Design Idea Sketches**
**Design:**



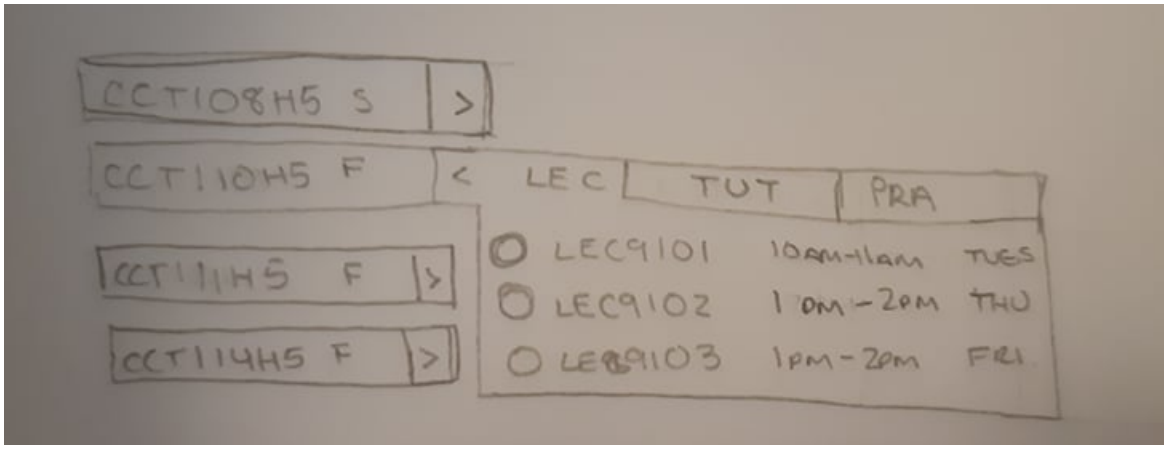

**Description:** The above design shows a different format of displaying courses in a way which is space efficient. Users can hover or click on course codes to expand them and select different lectures and tutorial times. Different sessions (fall, winter, and full session) are expressed through different colours; blue for fall, green for winter, and orange for full session. The idea was inspired by honeycombs.

**Design:**




**Description:** The above design shows another different format of displaying courses in a space efficient format. Much like the honeycomb design users can hover or click on course codes to expand and choose different lecture times. Different sessions are represented in different colours. The design was inspired by bubbles and bubble wrap.

**Design:**

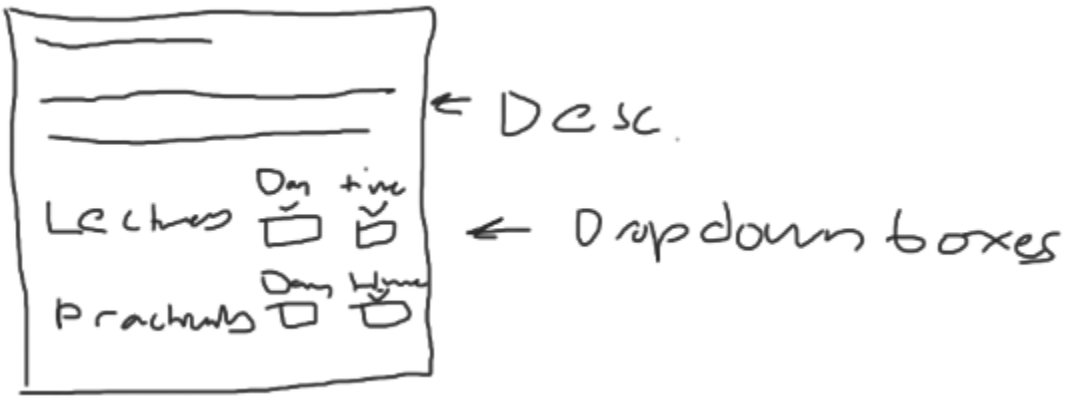



**Description:** The above design shows a different way of showing courses in a very compact and space efficient way. Course codes are listed and users can click a drop down to open different lecture or tutorial times (PRA, TUT, and LEC are shown in different tabs). The idea was inspired by sub folders and directories in File Explorer and the different tabs were inspired by browser tabs and sticky notes/bookmark tabs/tags.

**Design:**

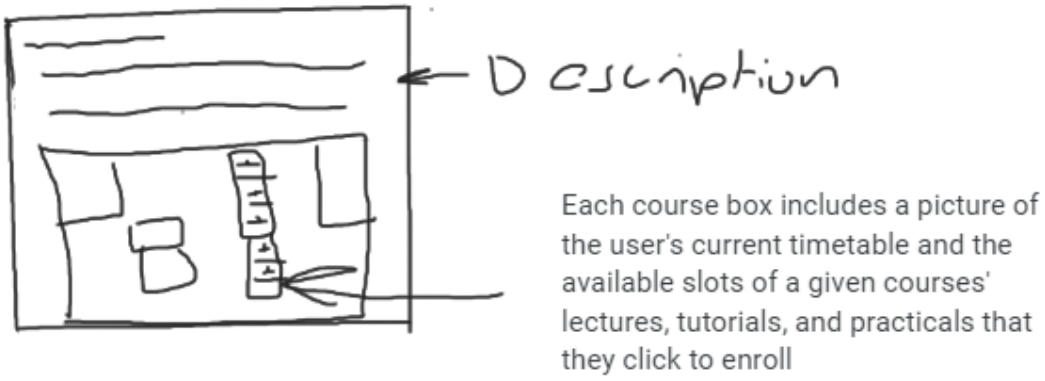

**Description:** The above design presents a different way of formatting the calendar/schedule with the list of courses to reduce scrolling. The little side schedule shows a mini version of the fall and winter schedule as users select their courses. Users can also select an option which shows you the schedule in full screen (like traditional timetable schedule). This provides immediate feedback to users as they choose their courses and they can get an idea of what their schedule will look like. This idea was inspired by the taskbar and split view on computers.

**Design:**



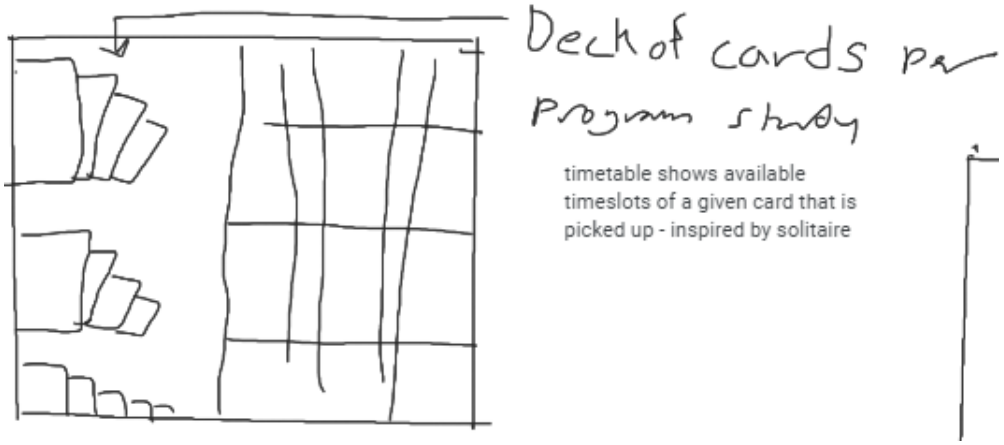

**Description:** The above design is inspired by scrolls that unravel. This is to offer more options to choose from without having to scroll down on the page. This is to be applied on the day and time options given that there are different days and time slots to choose from for one course.

**Design:**



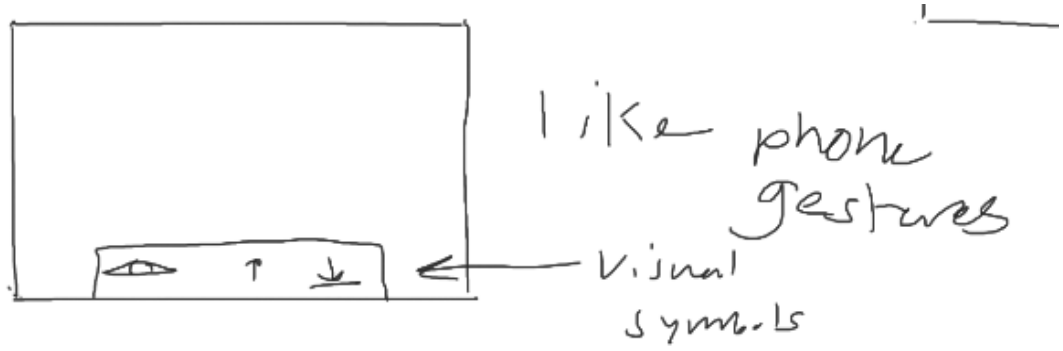

**Description:** The above design is inspired by a map, offering a visual focused interface based on the user's timetable. Each course box would then show the user's timetable and the available slots that course occupies for the user to select. This allows the user to get a better picture of what their schedule would be like and then adjust each course slot accordingly.

**Design:**



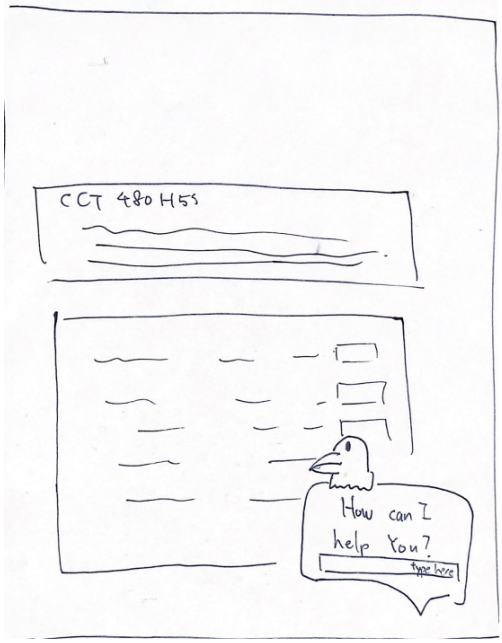

**Description:** The above design is inspired by a deck of cards. On the left side are piles of cards that represent each course box. The different piles signify the different programs of study that the user can swipe, shuffle, and select from. On the right, is the user's timetable where they place their cards/courses on. The timetable would show the different spots a course can be placed depending on the course day and time.

**Design:**



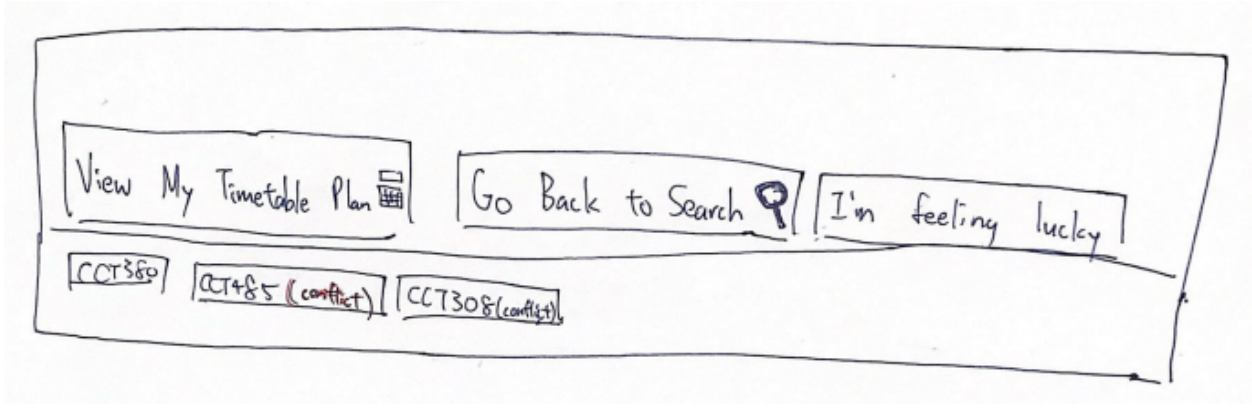

**Description:** The above design is inspired by the gesture bar found on android phones. This bar takes the current bottom bar of the timetable planner page and turns the functions into visual icons such as viewing the timetable, going to the top page, and downloading the timetable.

**Design:**



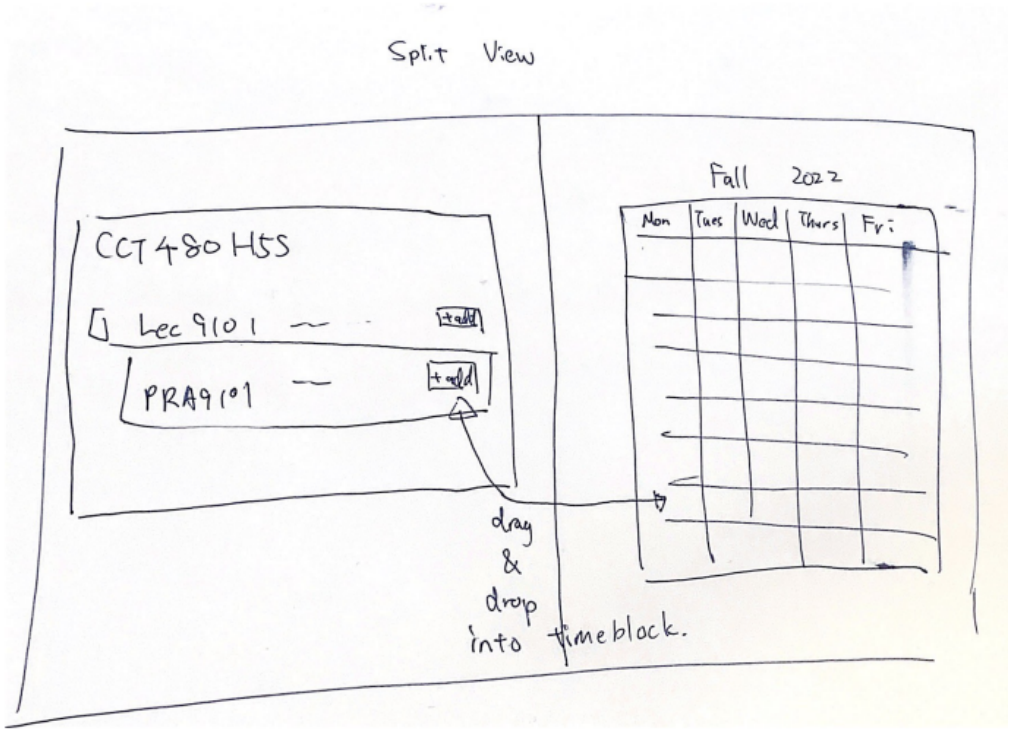

**Description:** The above design is inspired by the text assistant chat bots that are seen on certain websites. The current toolbar is limited to only 3 options and doesn't exactly cover a variety of functions. With the help of a smart bot as such, a lot of questions that students may have can be answered within seconds without having to call the office of registrar, thus making the planning process less stressful.

**Design:**



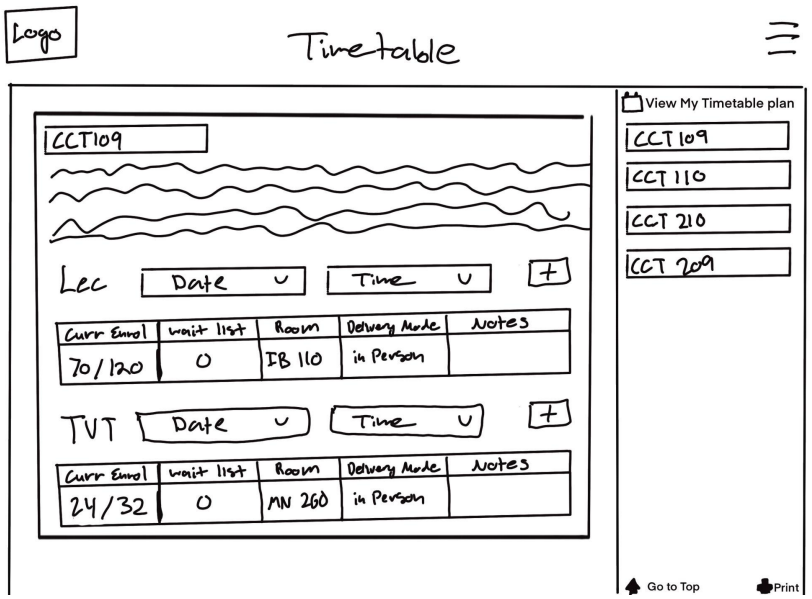

**Description:** The above design is a modified version of the current toolbar with better visibility in hopes of eliminating confusions with the current one. An additional "I'm feeling lucky" button inspired by the Google search page is also added, which suggests random electives for the user to enroll in, making the whole process a little more enjoyable.

**Design:**



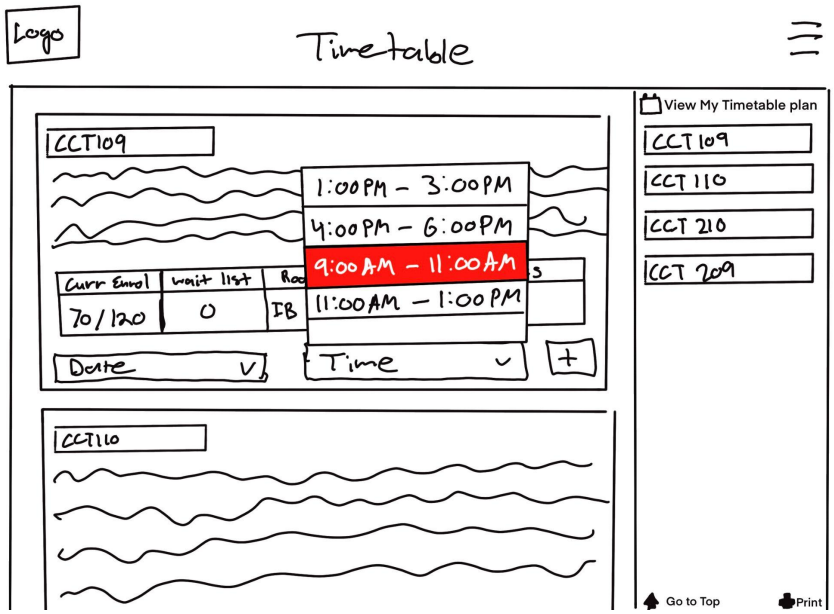

**Description:** The above design draws inspiration from a generic split screen view that's often seen on code editor programs where one can see the changes live on the right side of the screen. The user will have the ability to drag and drop a course from left onto a specific timeblock on the right or they can simply press "+add" and will see their results reflected right away on the right side of the screen. This feature aims to minimize the amount of time spent scrolling on the planner website.

**Design:**



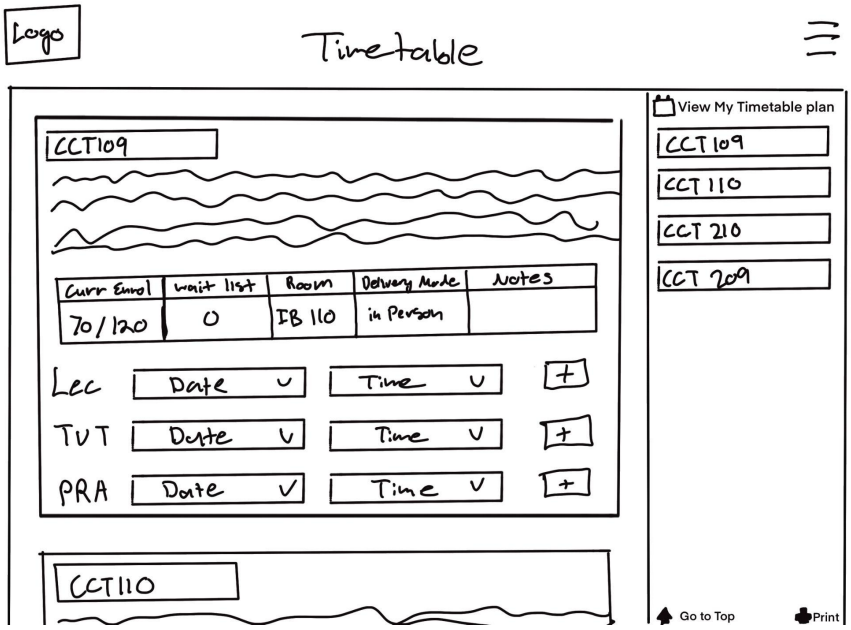

**Description:** The above design is a more simplified way in showing course info. Having a table on the left to explain essential information while having the user be able to pick date and time through a dropdown window made more sense and helped compact the information.Inspired by file and desk organizers, and cabinets. The page is clear and organized making it easier to understand.

**Design:**



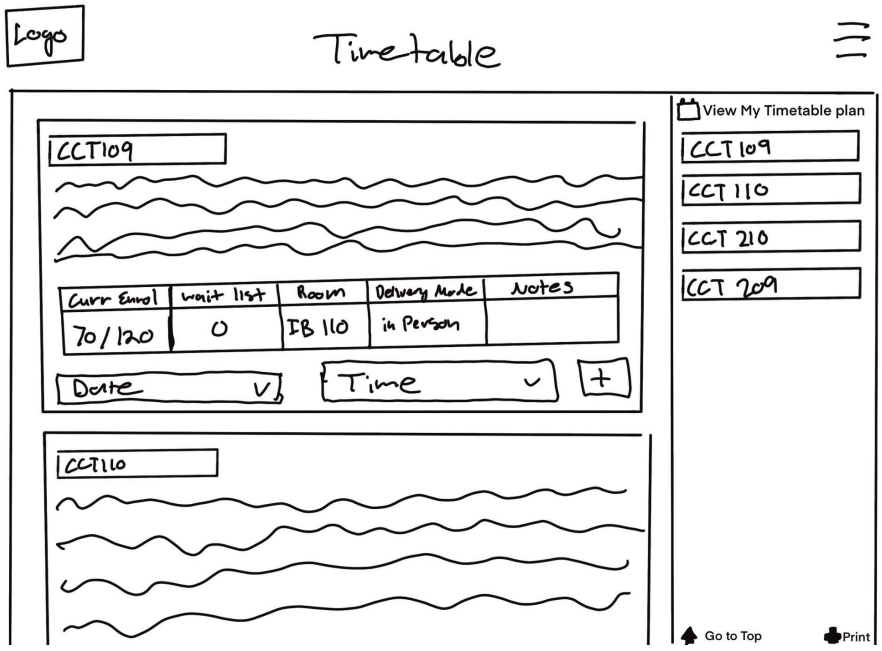

**Description:** The above design is when a user clicks on one of the dropdowns and has a conflict with one of the selected courses they already have in their schedule. It turns red and disables the option for the user to even place it on their schedule.

**Design:**



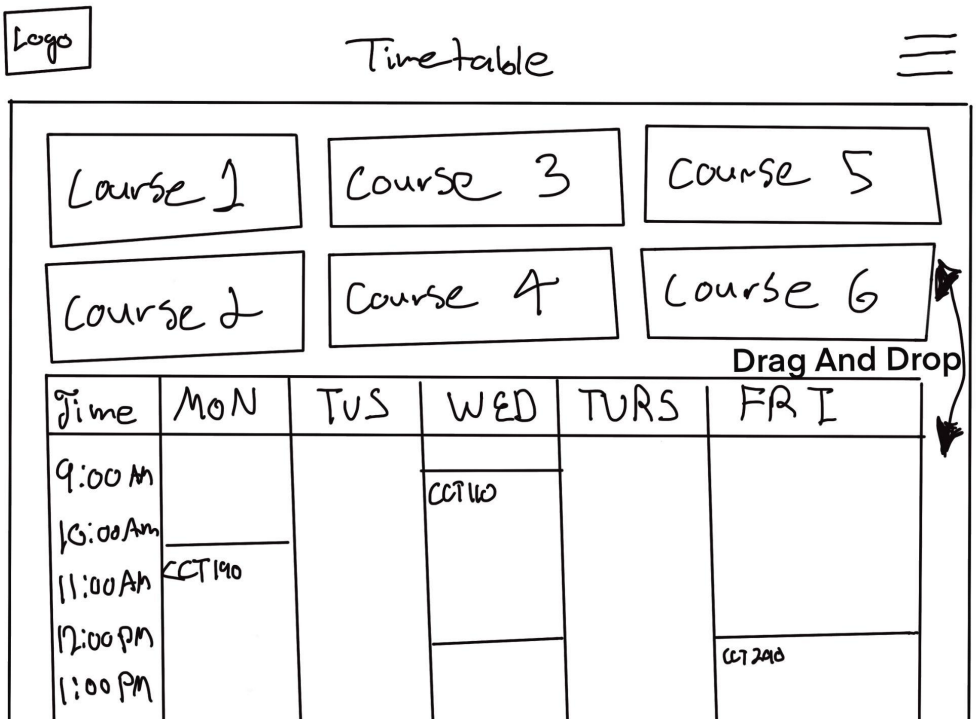

**Description:** The above design is a variation of the previous one with different dropdowns for Lecture, Tutorial, and Practical in the same class room.

**Design:**



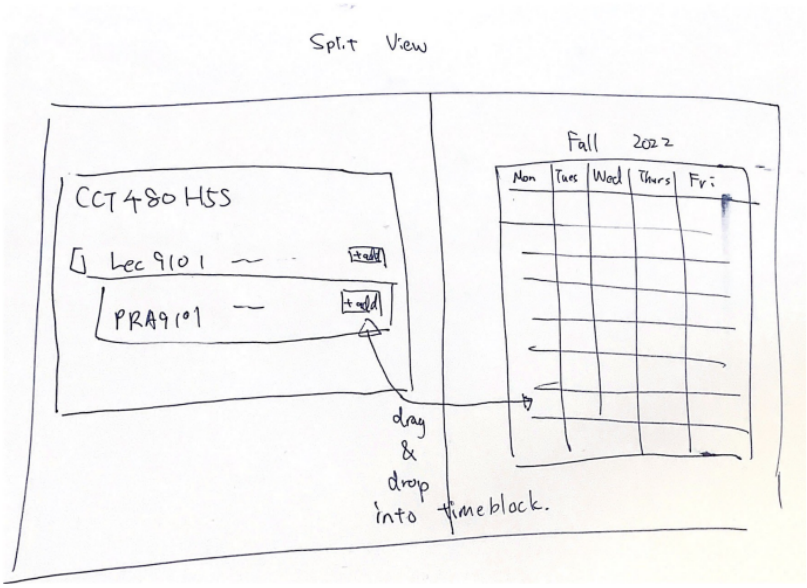

**Description:** The above design is the same as the previous one with a class that has only a lecture and no practical or tutorials.

**Design:**



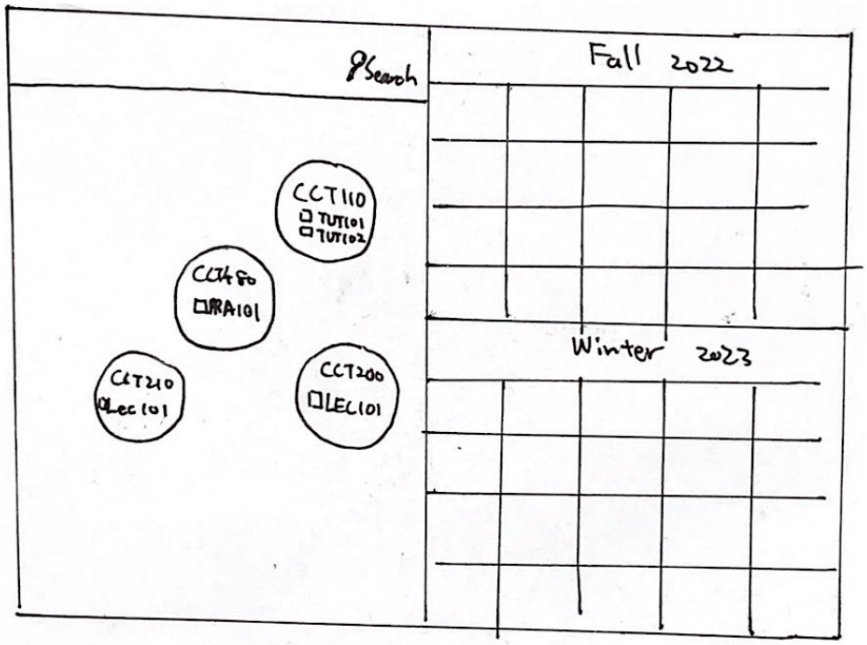

**Description:** The above design is an example of a split view design where the top half is courses and the bottom half is a calendar view. The user can drag and drop the courses from the top to the bottom to place it on their schedule. If there is an error or conflict, the user won't have the option to place it.

**Chosen Design From Part 1 of 10 + 10 Design**



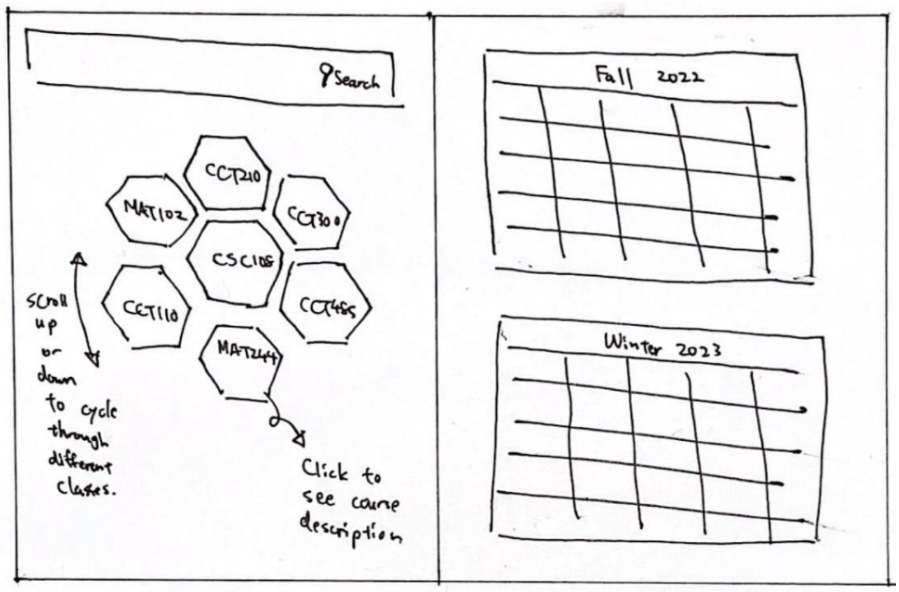
## Part 2 of 10 + 10 Design: Iteration and Variations of Chosen Design

**Design:**



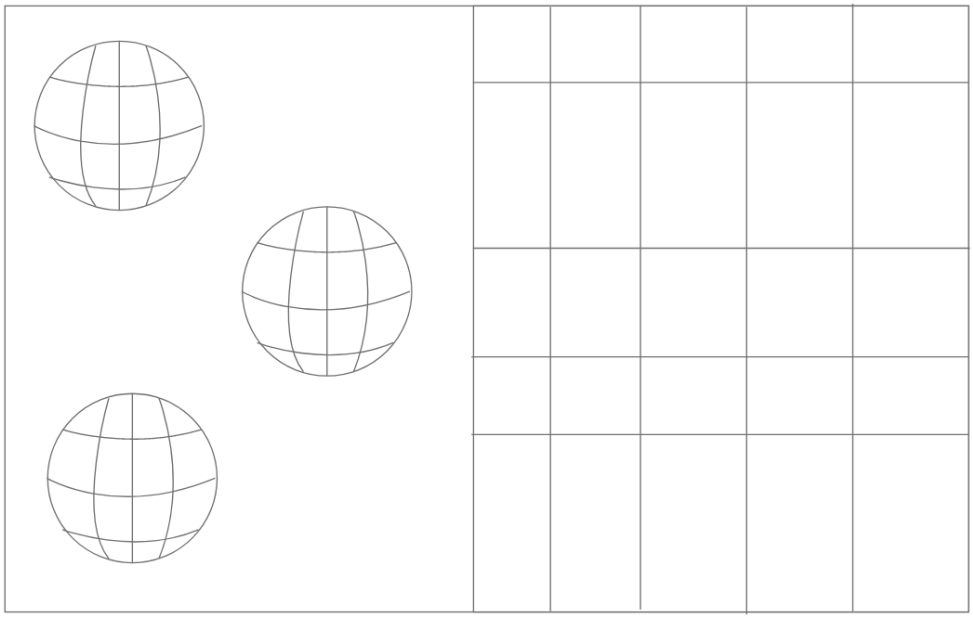

**Description:** The above design incorporates the bubble/bubble wrap idea that was first introduced in the first round ten ideations of the 10 plus 10 sketches with the split view feature to make it more interactive than generic split view screens. Additionally, to further honor the inspiration from bubble wraps, users will be able to double tap the bubbles to make them disappear, just like a real world bubble wrap.

**Design:**



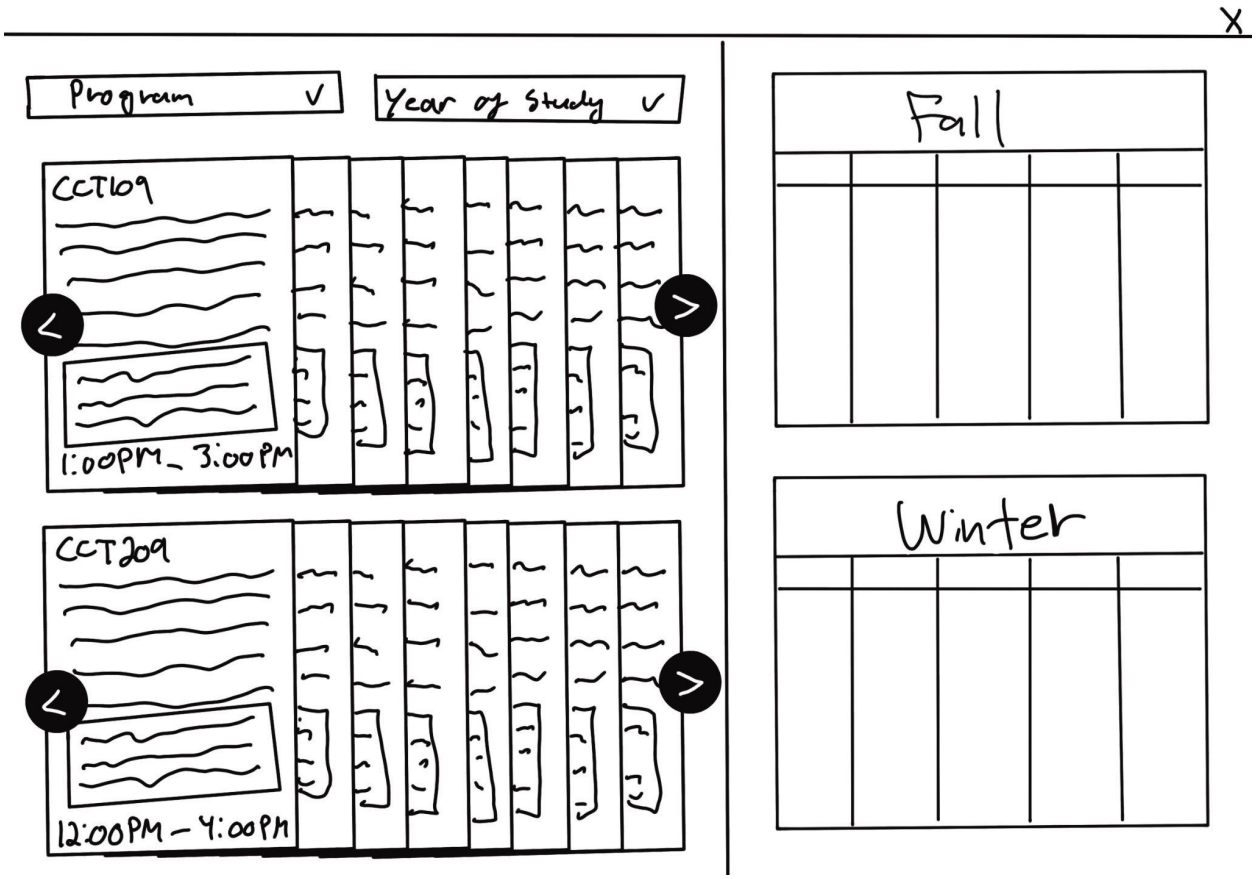
**Description:** The above design brings back the honeycomb hexagon shapes inspiration in the first half of the ideation process with some adjustments to reduce information overflow on a relatively limited space. When a user clicks on one of the hexagons, it will expand to reveal the course descriptions and lecture/tutorial sections. The user will also be able to scroll up/down or left/right to cycle through each chosen course.

**Design:**



**Description:**
This design is inspired by globes where each globe represents a program of study. The user can then move around the surface of the globes to look at different courses then select that course, drag, and drop it onto the timetable on the right.

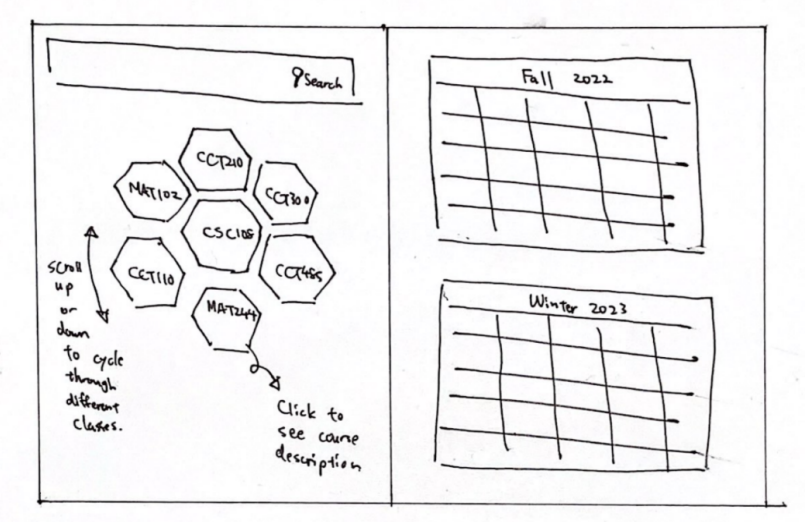
**Design:**




**Description:** The above design is influenced by the deck of cards idea that we had in the first round of the 10 plus 10 sketches design. We combined it with the split view to make a more fun and friendly design. Just like a deck of cards, the user can swipe and shuffle between them to see what course they want to add on their timetable. They can then drag and drop it on the right side of their calendar.

## Assignment 3: Project Proposal, Interface Design, and User Flow Map

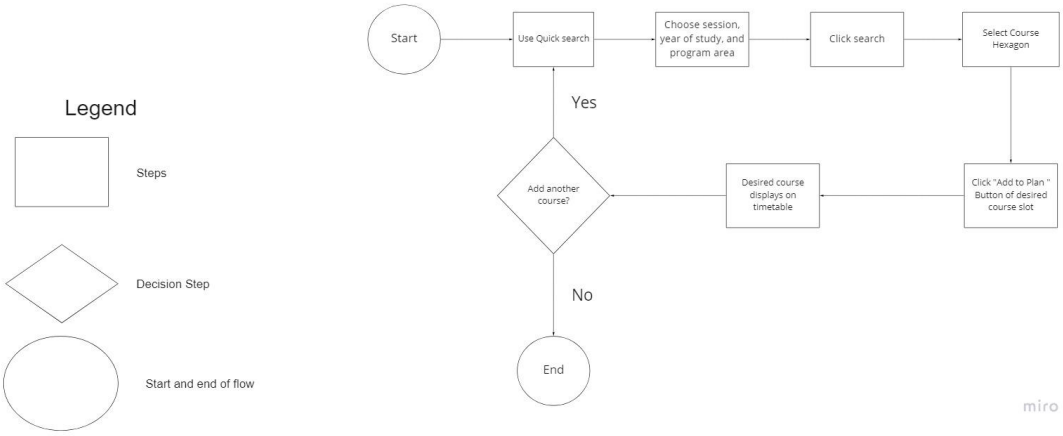
**Final Chosen Design from Part 2 of 10 + 10 Design:**




**User Flow**

**Changes from A3: We added a User Flow Map to reflect different paths users would take on our prototype.*

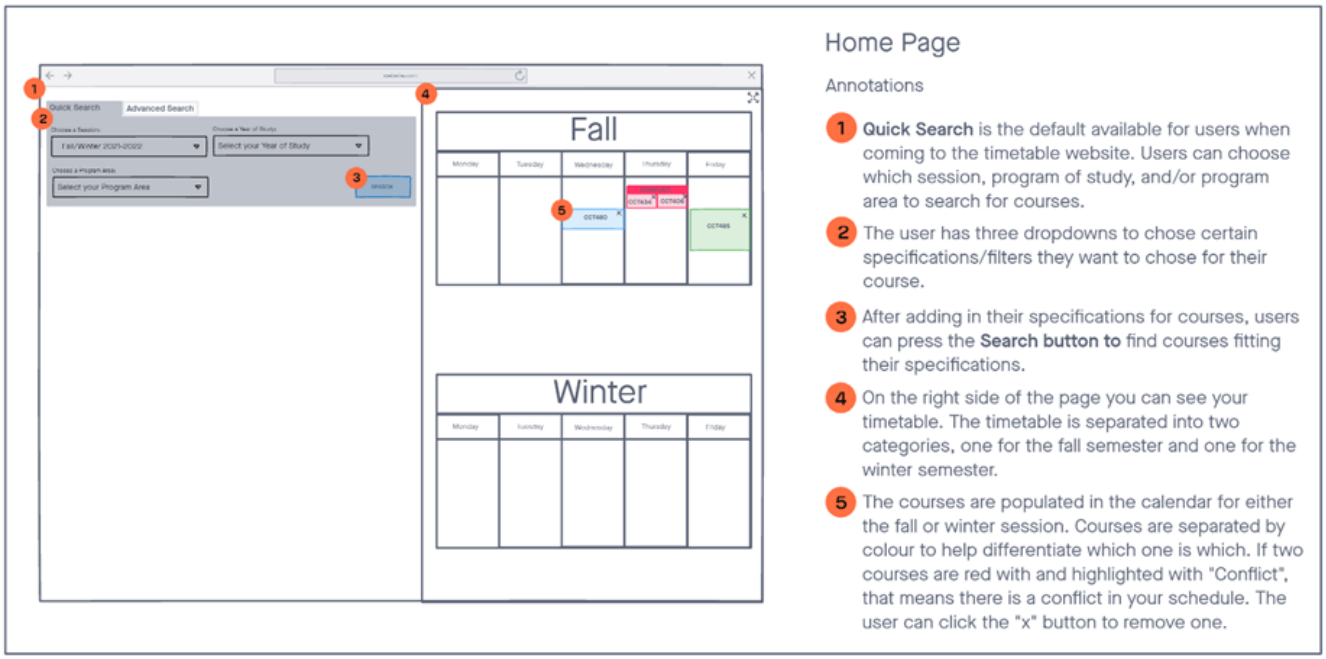
The user flow we chose to focus on is the flow of searching for and adding a course to your timetable (for the sake of our wireframe we chose 4th year CCT courses.

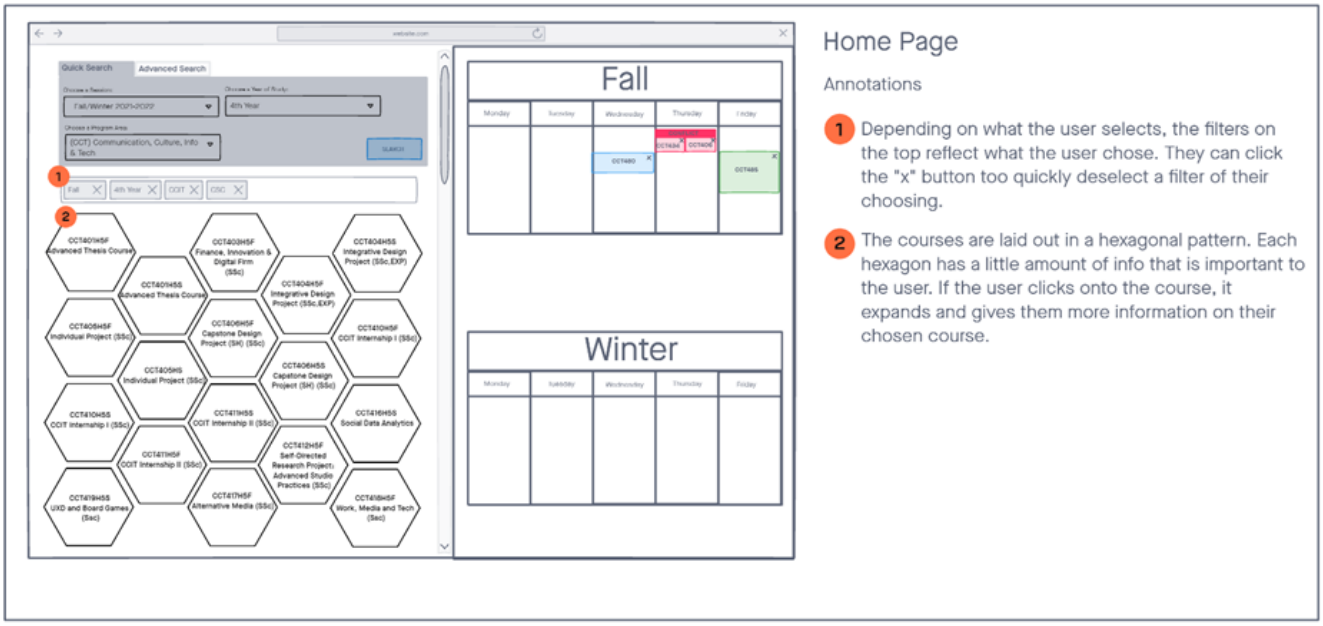


# Assignment 4: Preliminary Functional Prototype

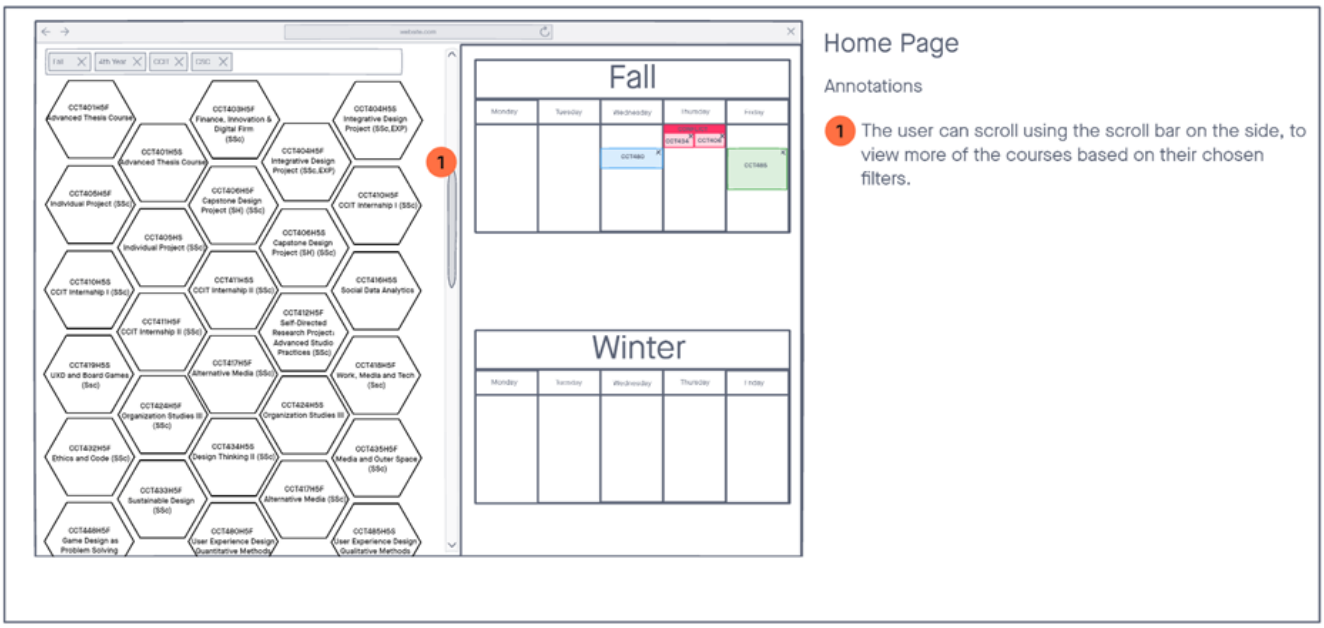
***Initial UTM Timetable Search Screen***

**Changes from A3: We added annotations to each page of the prototype to explain what different features/elements are for and what they do.*

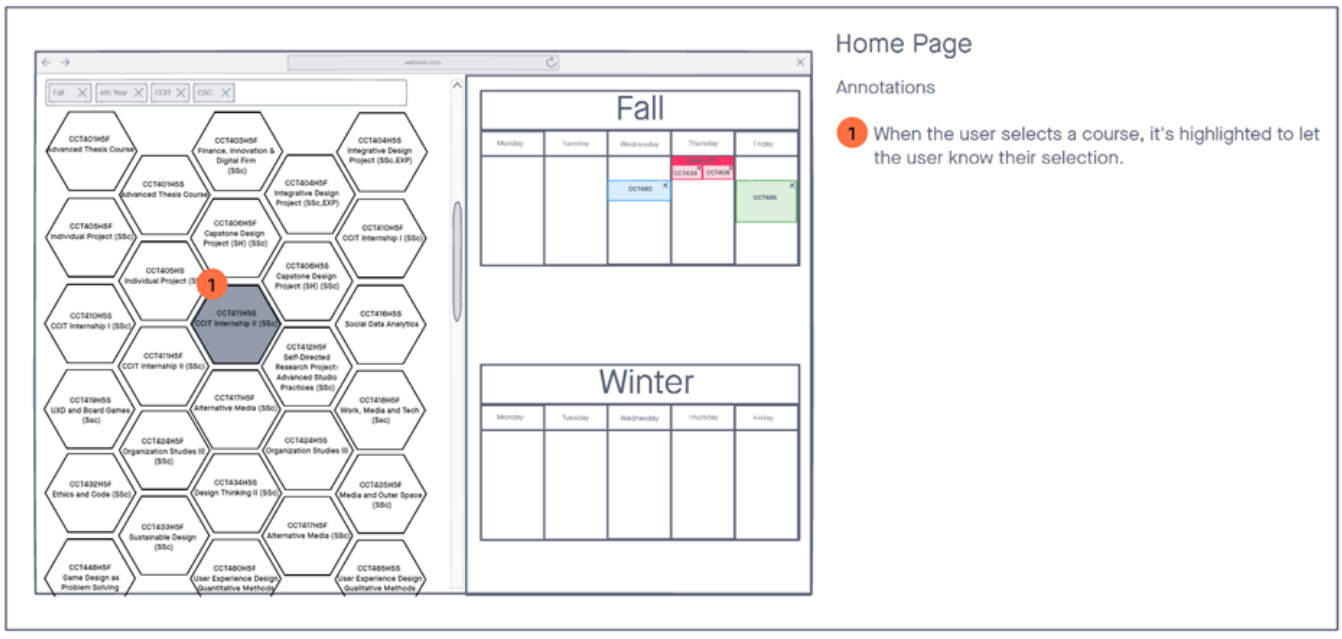


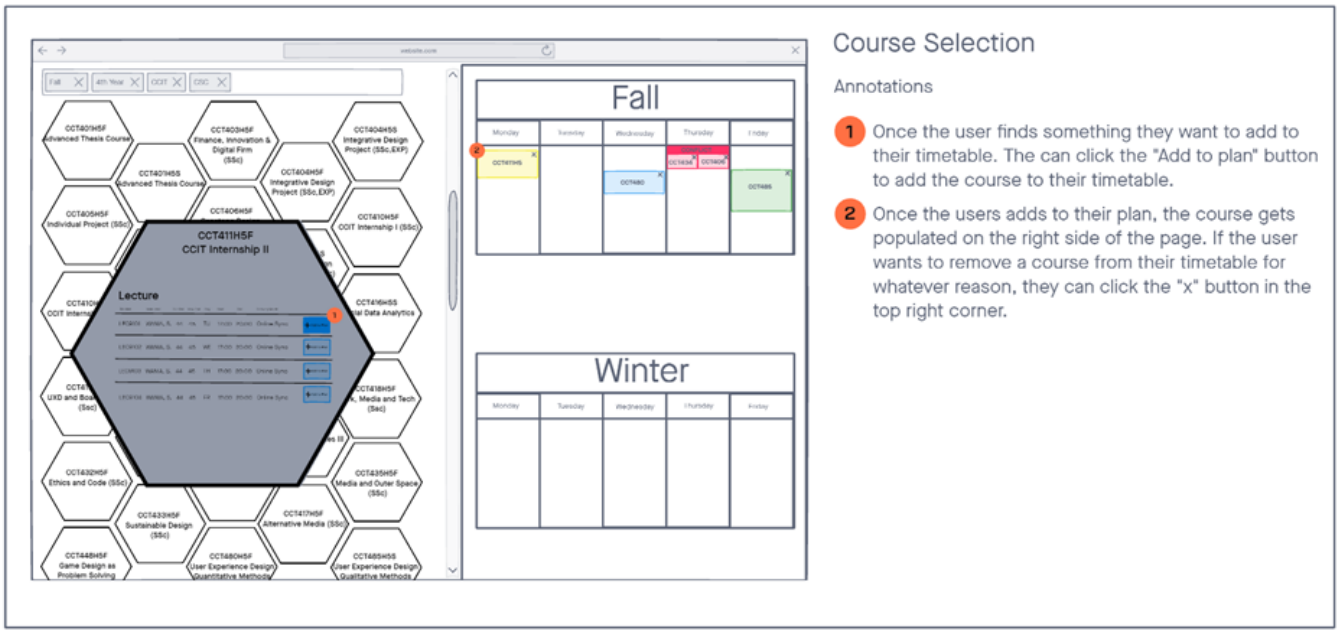



***Search Results for Fall/Winter 2022, 4th Year CCT Courses***

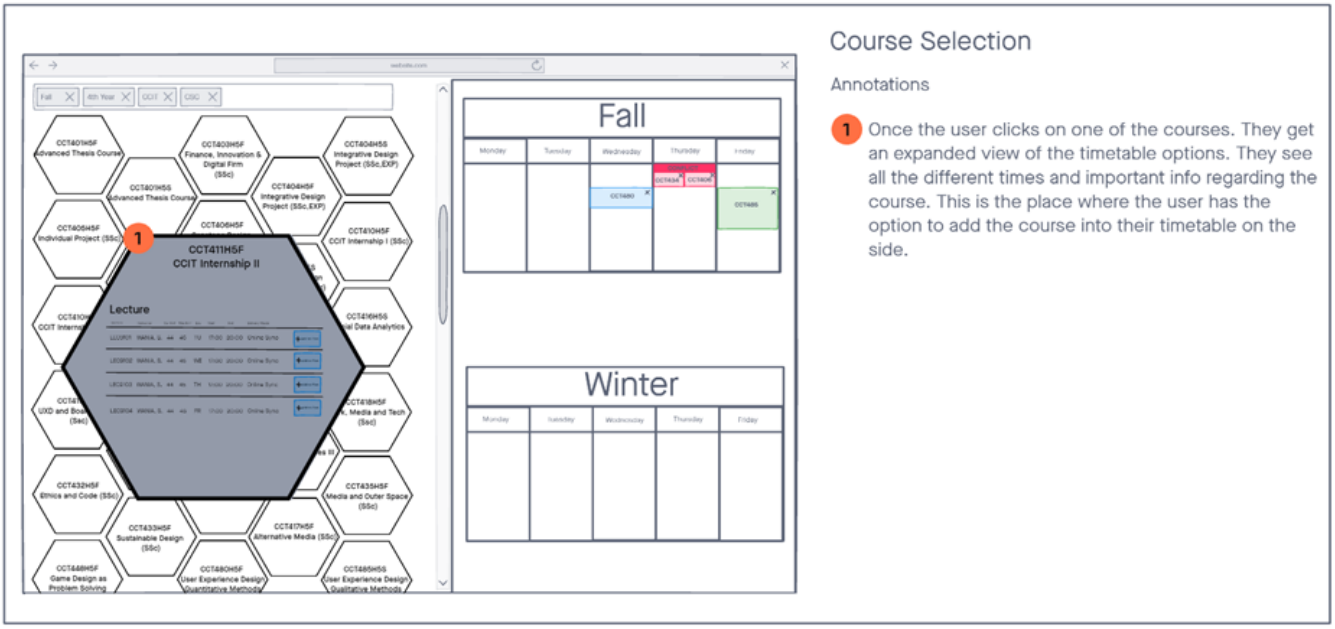



***Picking a Specific Course to Add to Plan***






***Adding Course to Plan***



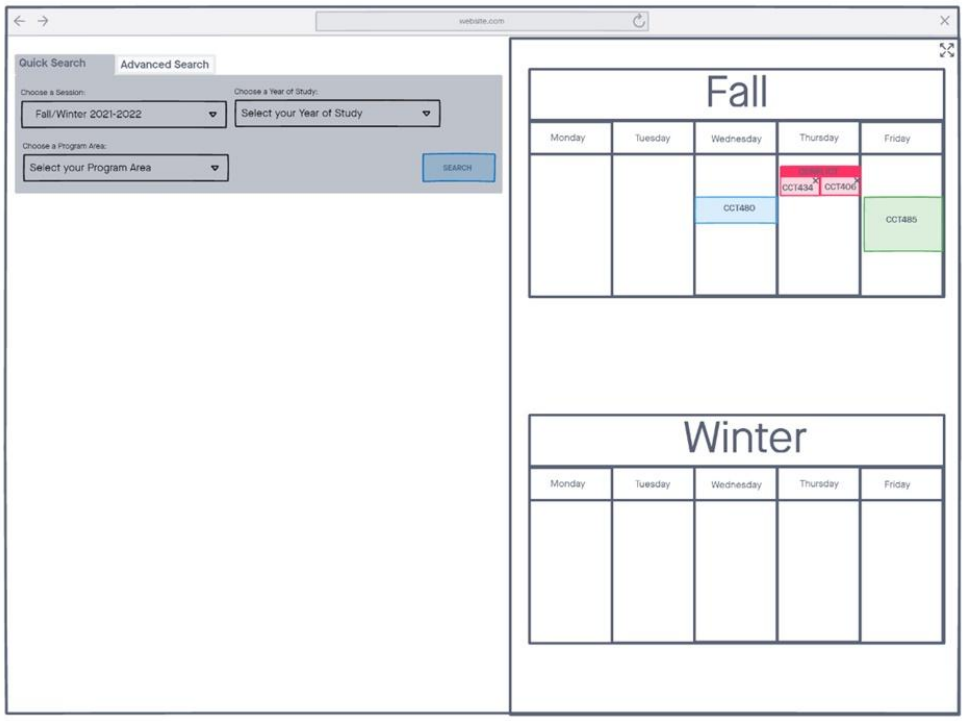

***Link to Wireframes:***

***https://manahilmasroor718356.invisionapp.com/freehand/Untitled-Qr5qUjyOT?dsid_h=bc77f8d5d0c43f1e87aef00217919c44bedc7a1e080a1ad33e01d1e3bd76e314&uid_h=accb26bbc3af0748df4a1546bdc88c3fb19b93bc99bc8156d935a0696a6d3e83***

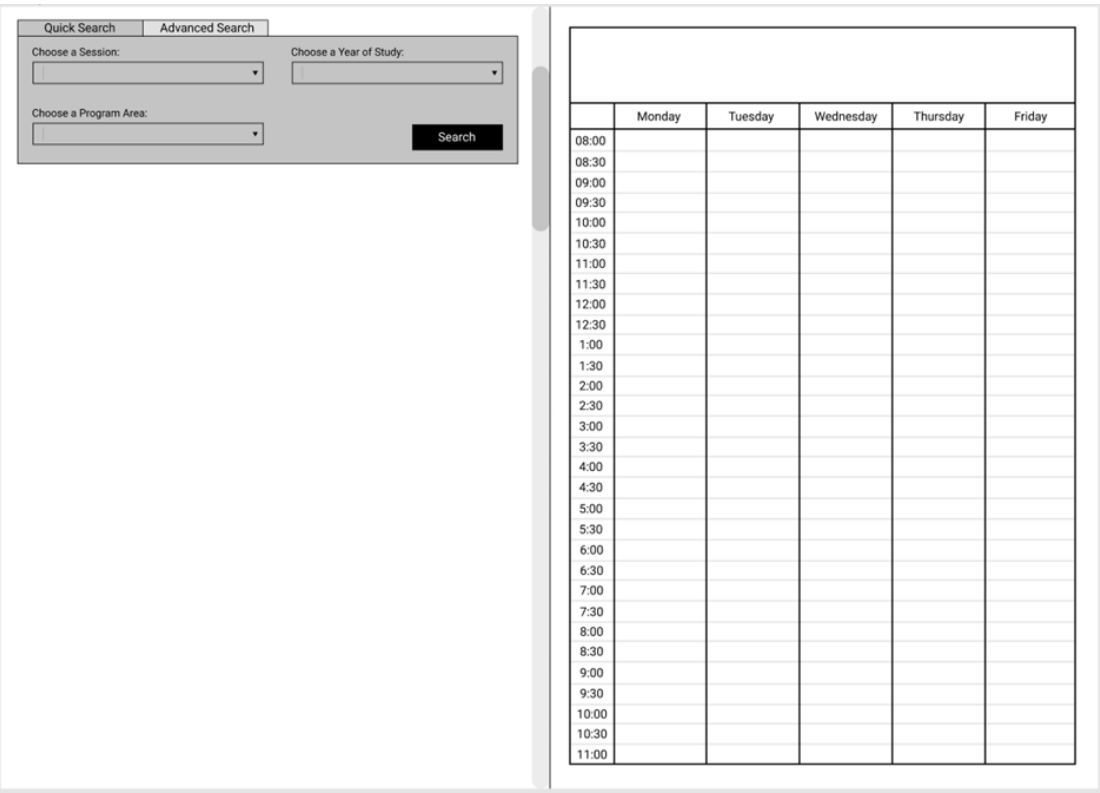
## Assignment 5: Guerilla Testing Results and Prototype Updates

**Prototype Updates and Description:**

***Old Initial UTM Timetable Search Screen Prototype:***



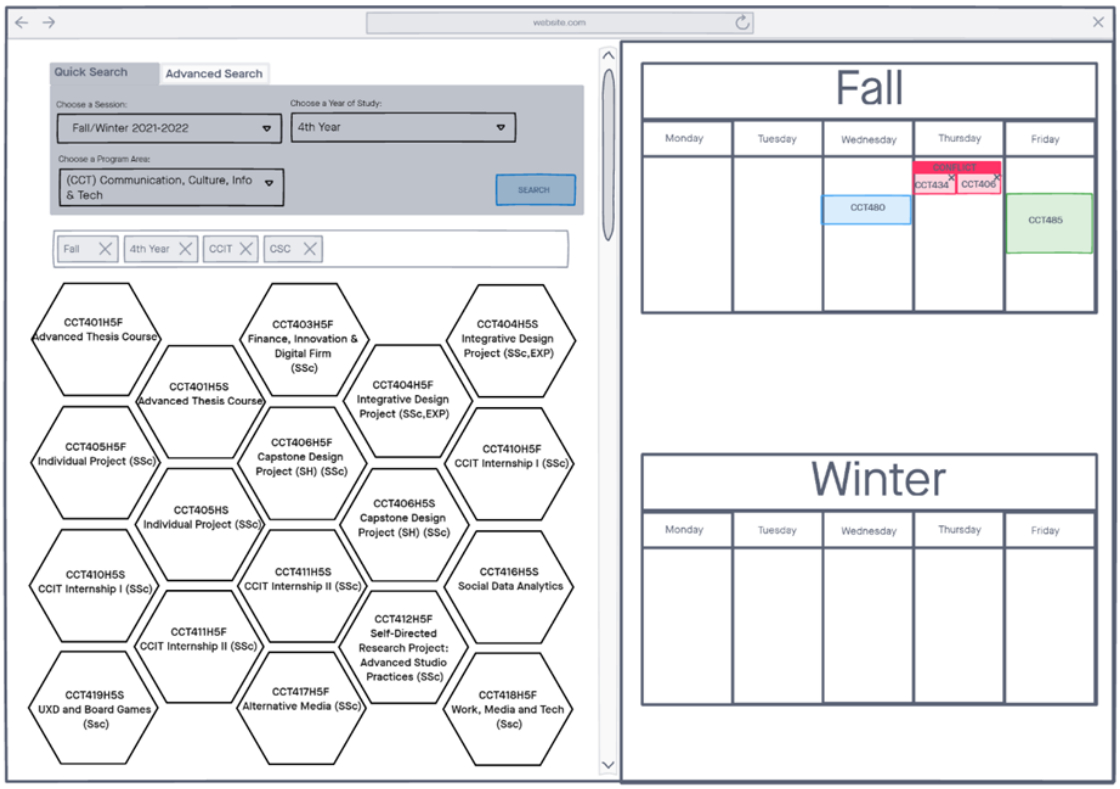
***New Initial UTM Timetable Search Screen Prototype:***



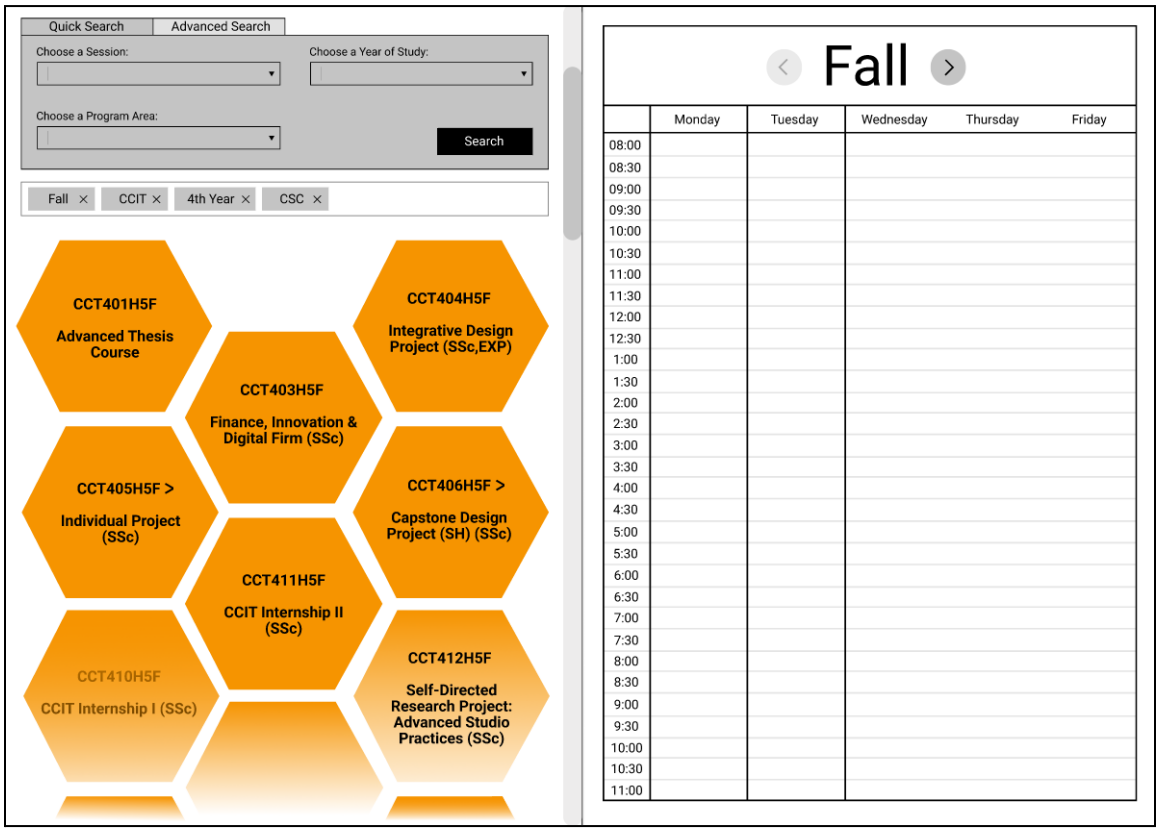

**Description:** The above shows the initial utm timetable planner website page. Users can search for courses on the left side of the split view via *Quick Search* or via *Advanced Search*. A blank timetable schedule is also displayed on the right hand side of the split screen.

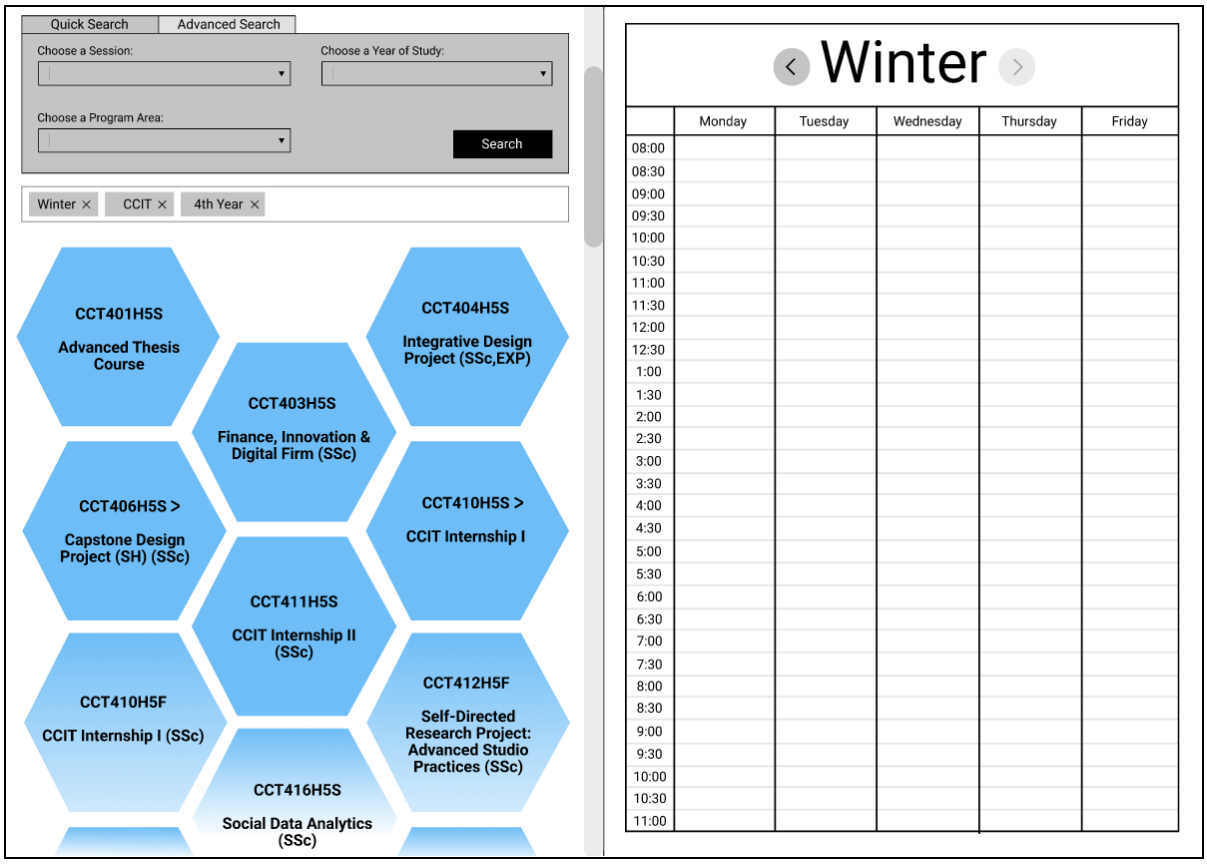
**Changes Made to Prototype:** As seen above, the new prototype is the same for the 'course search' side on the left. The right hand schedule now **only shows one semester (Fall or Winter or Summer) schedule** as users indicated the need to have the schedule be bigger and to utilize space more efficiently. User 1 from our testing expressed their desire for the 'Fall' or 'Winter' schedule titles to be bigger, while User 2 conveyed they would like to see the fall and winter schedules separately. User 3 & 4 also had to zoom in to even see the course codes in the schedule for the first version of our prototype. The new prototype includes the **time slots on the schedule** as User 4 & 5 suggesting adding times on the schedule to make it more readable.

***Old Search Results for Fall/Winter 2022, 4th Year CCT Courses Prototype:***




***New Search Results for Fall/Winter 2022, 4th Year Prototype:***





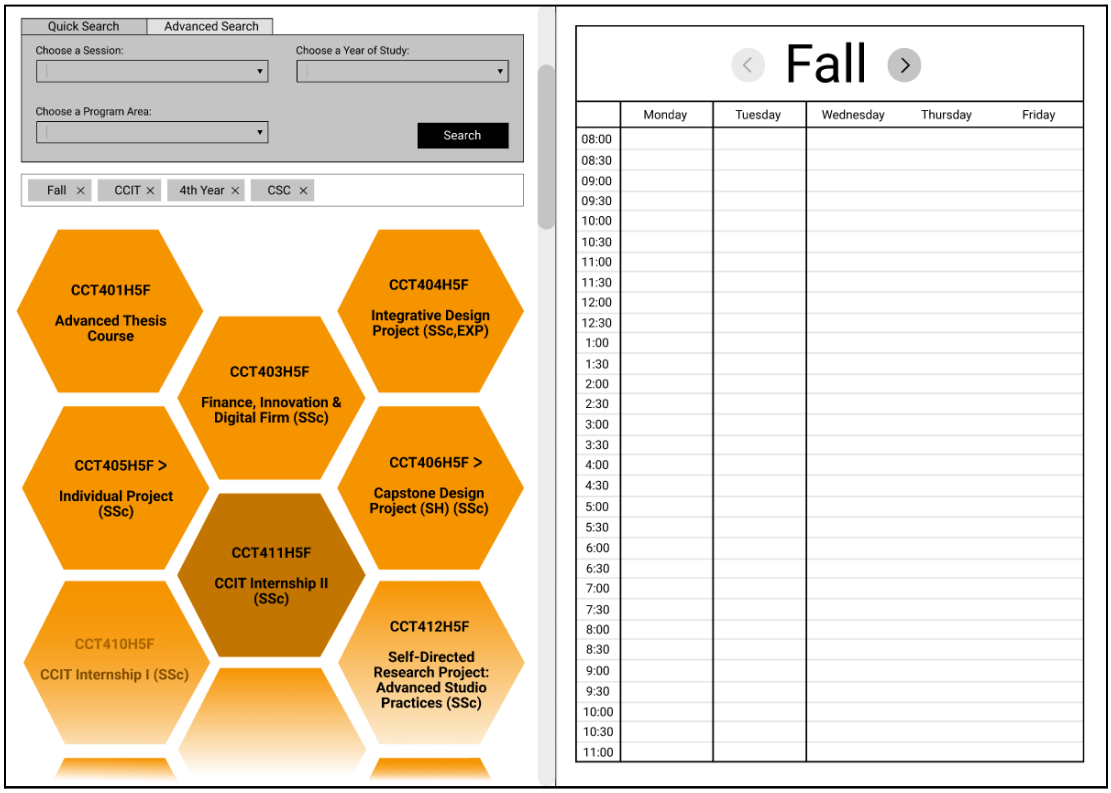

**Description:** The above prototype shows the screen users see for the results of their search; CCIT, Fall and (one for Winter), 4th Year. Each Hexagon displays a course code and course name. The courses in the above image are colour-coded with orange for fall or blue for winter. The right hand 'schedule' screen shows either a Fall or Winter semester schedule. The schedules are able to be changed from Fall to Winter (or vice versa) by arrows next to the semester session on the top of the schedule.

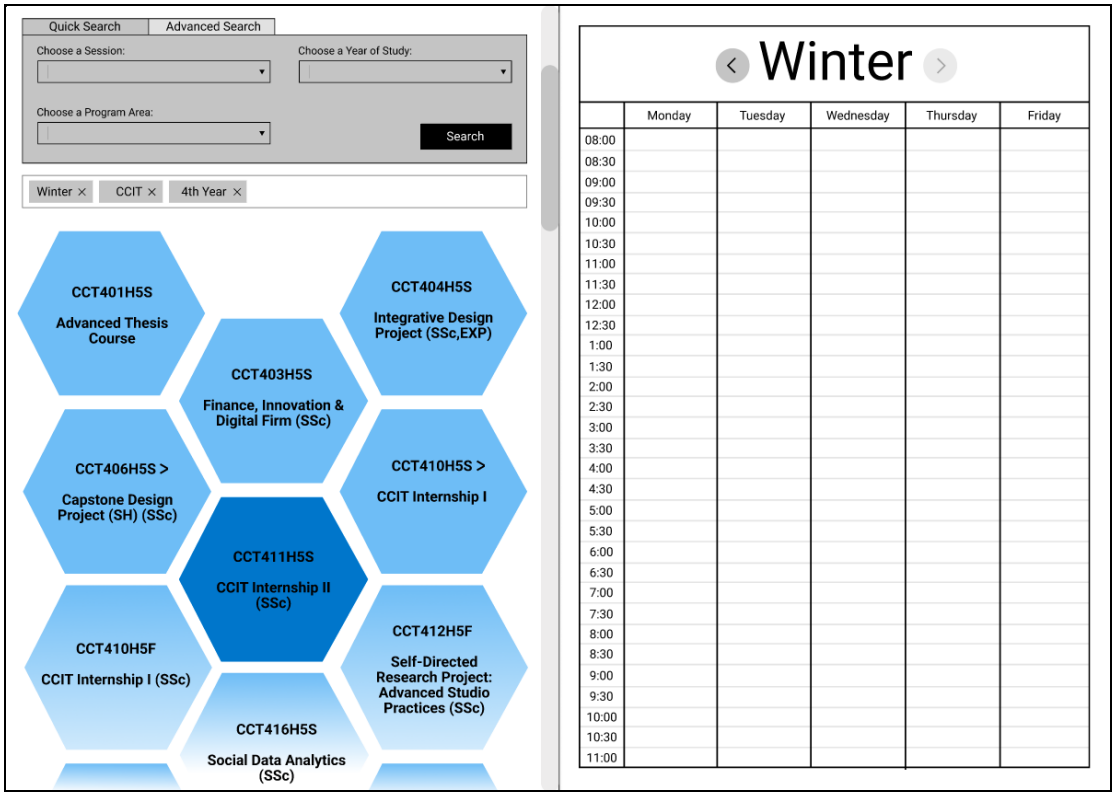
**Changes Made to Prototype:** The number of rows of **hexagons has decreased** from 5 to 3 in our new iteration, allowing for bigger hexagons and **better visibility for our course code and name**, as well as **less cluttering and more readability** for users. The design decision was made as User 1, 2 conveyed they believed that the course code text was too small and Users 1, 2, and 3 indicated trouble with following along the pattern to find the course and believed the search results screen to be too overwhelming. The new prototype 'course search' side features **colour-coded courses**; orange for Fall and blue for Winter. Users 2, 4, 5 conveyed their frustration with not knowing which course was Fall or Winter without looking at the course code which they also exclaimed was too small to read. The new prototype also allows changeability between **viewing fall or winter schedules** as Users 1, 2, 3, & 4 expressed their desire to separate the fall and winter schedules for our first prototype iteration (also referenced on initial utm timetable screen).

***Old Hovering Over Desired Course Prototype:***

N/A. The old iteration did not include a hover state of desired courses.

***New Hovering Over Desired Course Prototype:***

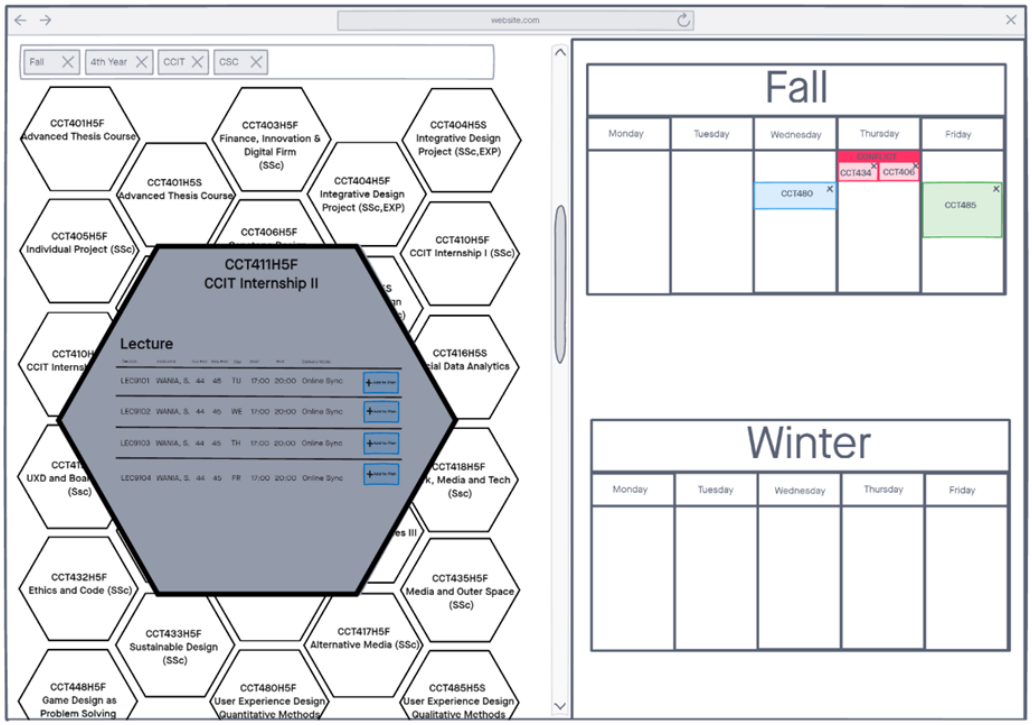


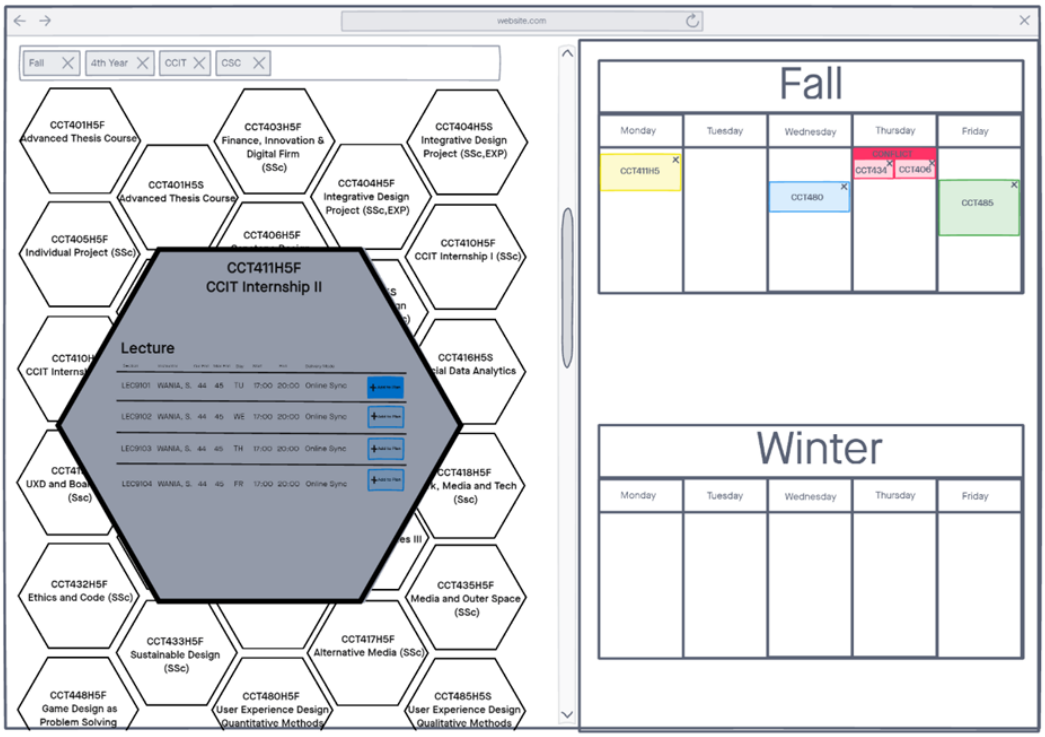



**Description:** The above image shows a user hovering over a desired course.

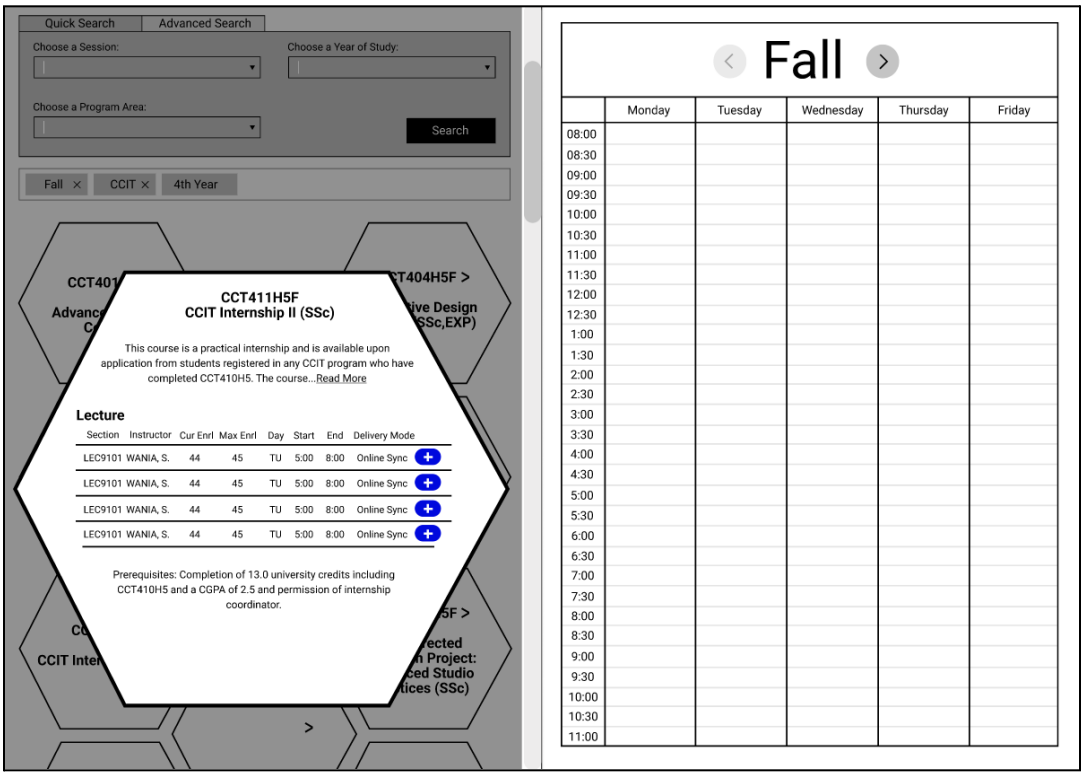
**Changes Made to Prototype:** The above image shows the **hover state** of the selected course to indicate that the course is clickable as users 1, 3, and 5 mentioned that their first intuition when they saw the hexagons was to drag and drop.

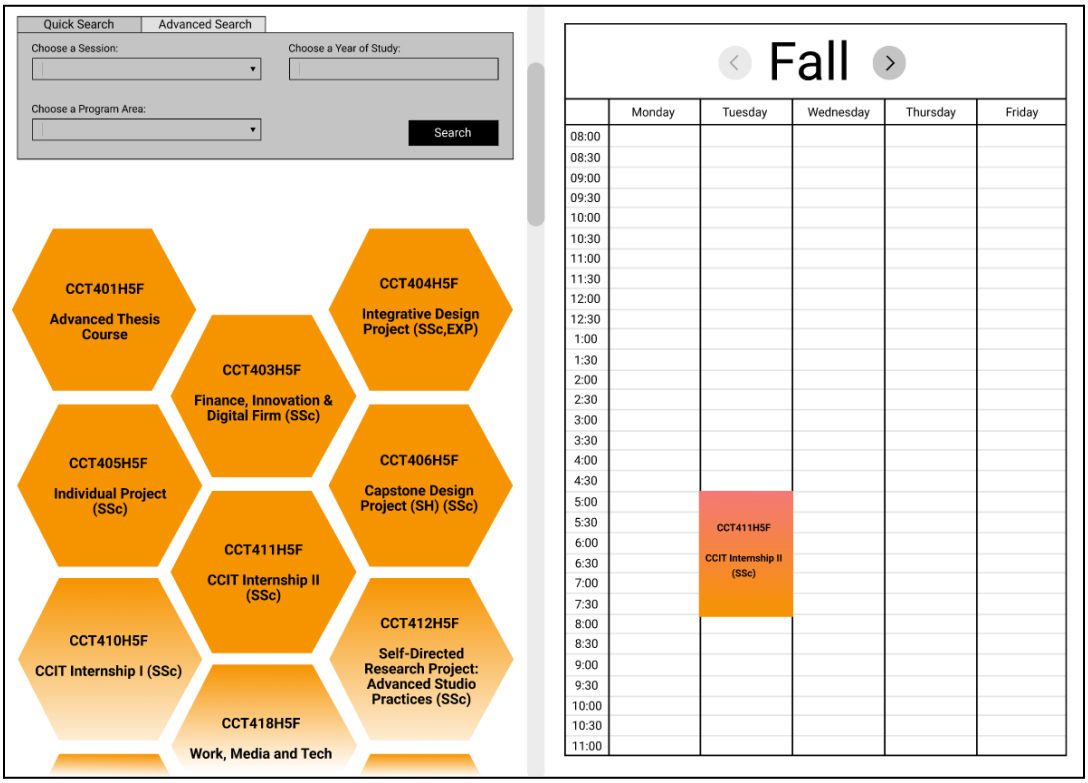
***Old Adding a Course to Timetable Schedule Prototype:***

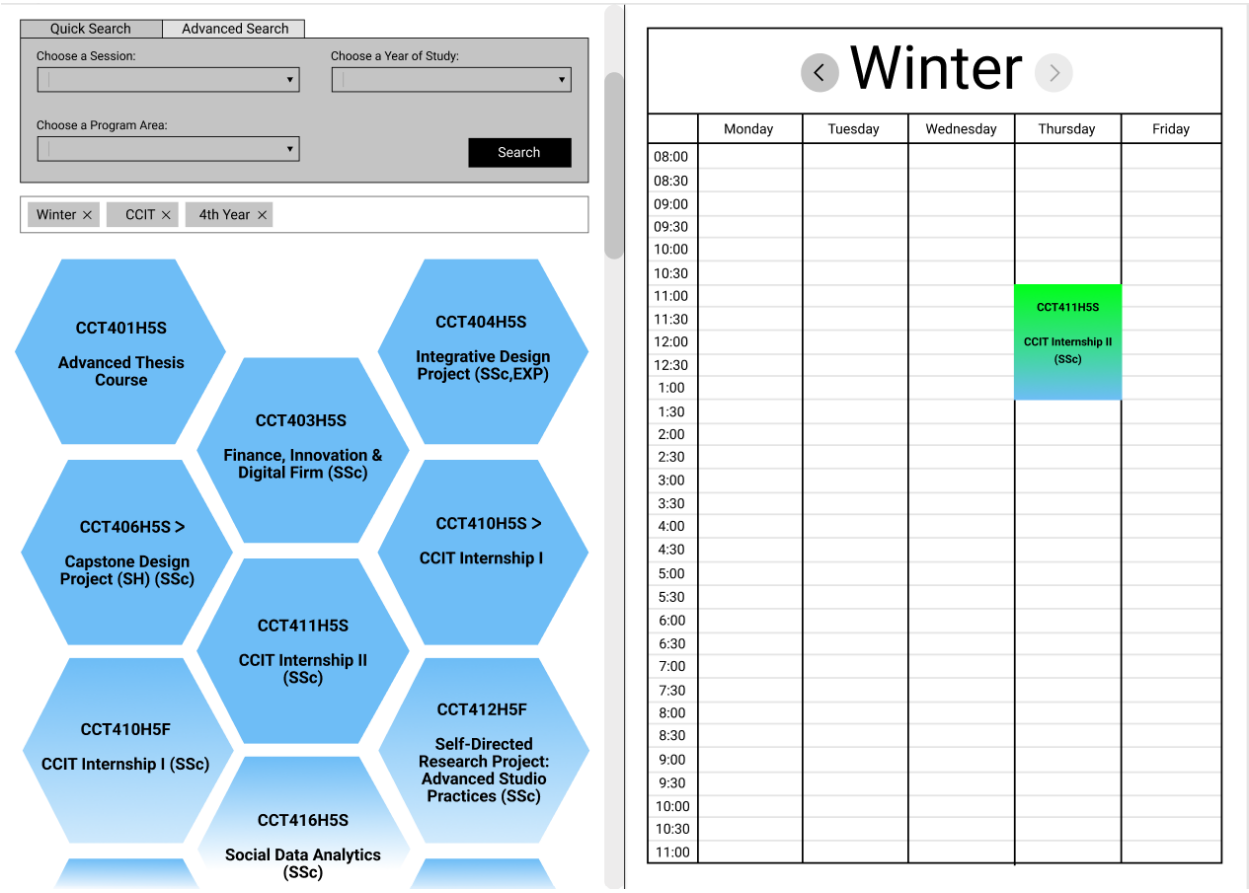


***New Adding a Course to Timetable SchedulePrototype:***






**Description:** The prototype shows the selecting and adding course to the timetable. Users are able to see different lecture timings/dates, assigned professors, amount of currently enrolled students and other information such as course description and prerequisites needed. Users are able to add the course to the timetable by clicking on the '+' button next to each associated lecture timing.

**Changes Made to Prototype:** The selected courses, in our new iteration, display both **course description and prerequisites** as User 4 suggested that course information be included.

**Possible Interactions to Benchmark:**

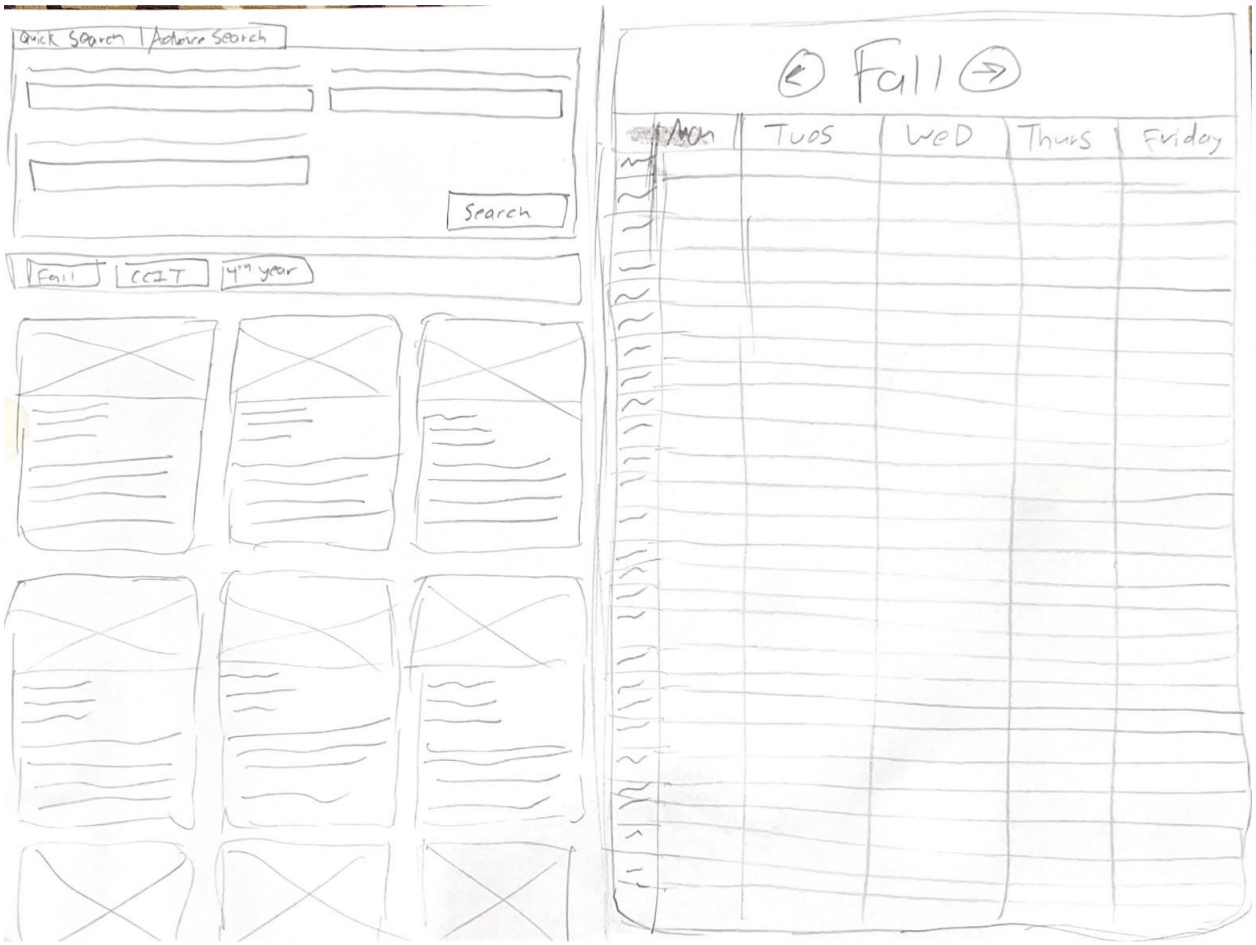
- An interaction we would benchmark is measuring how many users interact with the hexagon buttons successfully by accomplishing the task of successfully finding and adding a specific course time. This is measured with a table to check off if they successfully accomplished each task involved with adding a course (searching, locating course, opening lecture times by clicking on course hexagon, adding the course to the timetable by pressing the '+' button next to the certain lecture session). We will be doing this by breaking down the task and asking them in task such as 'Please show me how you would search for a 4th Year CCT Fall Course', 'Please show me how you would see the

different lecture times for CCT411', 'Please show me how you would add LEC9101 to your fall timetable' etc.

- Another interaction we would like to benchmark is measuring how many users click on the desired course instead of trying to drag and drop it. We will be doing this by having tables with users and then columns that specify if they tried to drag/drop, click or 'other'. It will also have a section for comments (if they tried to drag and drop and then clicked or any additional comments/observations).
- We would also like to benchmark how many users clicked on the arrows to switch between Fall and Winter timetables to measure if it was intuitive enough for them. This will be done by asking them as a task how they would see their Winter timetable schedule if they were currently in Fall schedule. We will measure this through time taken and amount of clicks it takes them (should be one click - since its one button). We will also make note of any other comments ex. If they click anything else.
- We will also be measuring overall satisfaction and how easy or hard the mental load was through a scale of 1 to 5. Ex. Overall how difficult was it for you to accomplish this task (1 - super easy, 5 - super hard). We will also be asking them their overall satisfaction with the task they just accomplished.

** We incorporated feedback given to us about A5, such as specified measurements for our benchmark tests and broke down tasks to be steps so we could measure those rather than if users were 'using the TimeTable as intended'.*

## Sketch of Alternative Design:




**Description:**

The sketch replaces the hexagon panels with cards instead. The cards help the user understand which course is going in chronological order and gives the user more initial information about each course. There is an image that can help the user visually get information about the course as well.

**Changed from A5: We changed our alternative design from the bubble alternative because we would have run into similar problems of readability and wanted to address the readability issue.*

## Assignment 6: Benchmark Testing Results & Prototype Updates

**Link to New Prototype:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/Untitled?node-id=25%3A139&scaling=scale-down&page-id=0%3A1&starting-point-node-id=25%3A139

**Link To Alternative Design:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/CCT480-UTM-Timetable-Planner?node-id=264%3A157&scaling=scale-down&page-id=5%3A3&starting-point-node-id=264%3A409

**Tasks Asked & Data Types Measured:**
**Changed from previous A6: We added tasks, what measurements we are taking other than satisfaction rate as we could not solely rely on satisfaction rate for our designs usability.*

- **Task 1:** Adding CCT411 to Fall Schedule
  - **(Onboarding) Task 1a:** Could you show us how you would search for Fall, 4th Year CCIT Courses
    **Changed from original A6: we originally had been searching for the fall 4th yr, CCT course as our first task, and measured it through likert scale. In our other iterations we changed this to an onboarding task and removed the likert scale as this quick search is the same in our design compared to the original UTM timetable planner and did not give us helpful results to determine the usability of our design.*
  - **Task 1b:** Could you show us how you would **see the lecture times for CCT411**
    - To measure this, we measure whether the user clicked on the hexagons to see lecture times or if they tried to drag and drop the hexagons. This is an example of **nominal** data types.
      **Changed from previous assignment 6: we added the measurement of if users clicked or dragged + dropped when opening up lecture times. Previously we only measured likert scale for this task which is not enough to give us data on if users interacted with the design as intended (click on hex).*
    - To measure this, we also used a likert scale to show their satisfaction on the process of selecting and finding courses. This is an example of **ordinal** data types.
  - **Task 1c:** Could you show us how you would **add a CCT411 lecture to your Fall Schedule?**
    - To measure this, we measured the number of clicks it took for them to add the course to the plan. We wanted to see how many users interacted with adding courses the way we designed them to, as it only takes one click (the blue '+' button next to each lecture section), any more than 1 click

needed would signify that the design was not intuitive enough for users and required change. We did this to measure for accuracy (more than 1 click, would mean #-1 = number of errors). This is an example of **interval** data types (not ratio as they need at least 1 click to do the task, 0 would mean they didn't do the task but all users' did).
**Changed from previous assignment 6: we added the measurement of how many clicks it took to add a course to the plan. Previously we only measured likert scale for this task which is not enough to give us data on if users interacted with the design as intended. This way we can see if there are any numbers of errors during this task of adding course (anything above 1 click is an error.*

- To measure this, we also used a likert scale to show their satisfaction on the process of adding the course to plan. This is an example of **ordinal** data types.

- **Task 2:** Could you show us how you would see you Winter Schedule
  - To measure this, we counted the number of clicks. As it was only one click (the arrow next to the Fall schedule title), we believed any more than one click to get the Winter Schedule would show that it is not intuitive enough to users and changes should be made. We did this to measure for accuracy (more than 1 click, would mean #-1 = number of errors). This is an example of **interval** data types (not ratio as they need at least 1 click to do the task, 0 would mean they didn't do the task but all users' did).
    **Changed from previous assignment 6: we added the measurement of how many clicks it took to switch to the winter schedule. Previously we only measured likert scale for this task which is not enough to give us data on if users interacted with the design as intended. This way we can see if there are any numbers of errors during this task to switch to the winter schedule course (anything above 1 click is an error).*
  - To measure this, we also used a likert scale to show their satisfaction on the process of finding the Winter Schedule. This is an example of **ordinal** data types.

**BenchMark Testing Data Collected:**
**Added testing results*

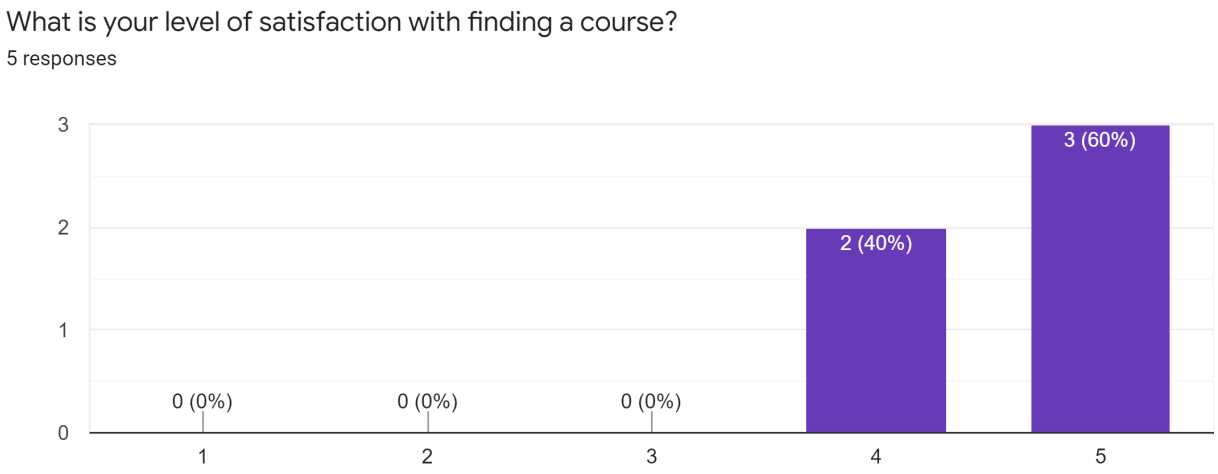
****Task 1A is Onboarding task so No Likert Scale****

***Task 1B Choice of Selection/Errors***

| *TASK 1B: Interaction with Seeing Lecture Times Hexagons* | | | |
|---|---|---|---|
| **Participant #** | **Click (Desired)** | **Drag and Drop (error)** | **Other (Specify)** |

| | | | [Error] |
|---|---|---|---|
| 1 | Yes | | |
| 2 | Yes | | |
| 3 | | Yes, Instinctively would like to click,drag and drop | Goes to advanced search and wants to scroll - instinctively |
| 4 | Yes | | |
| 5 | Yes | | |

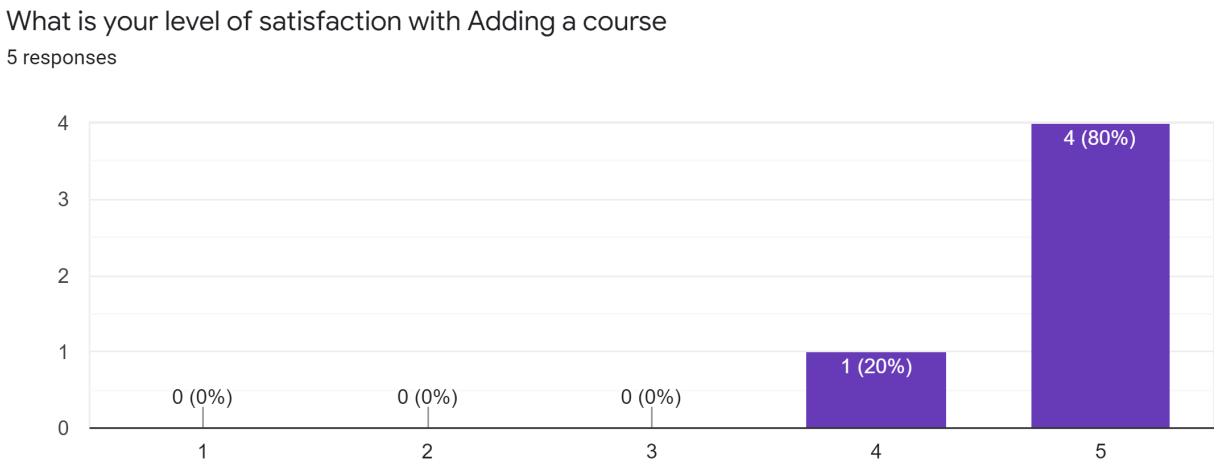
***Task 1B Likert Scale:***




***Task 1C Number of Clicks***

| Participant interacts with the Timetable as Designed (Accuracy) | |
|---|---|
| Participant | TASK 1C: Number of Clicks to Add to Plan |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |

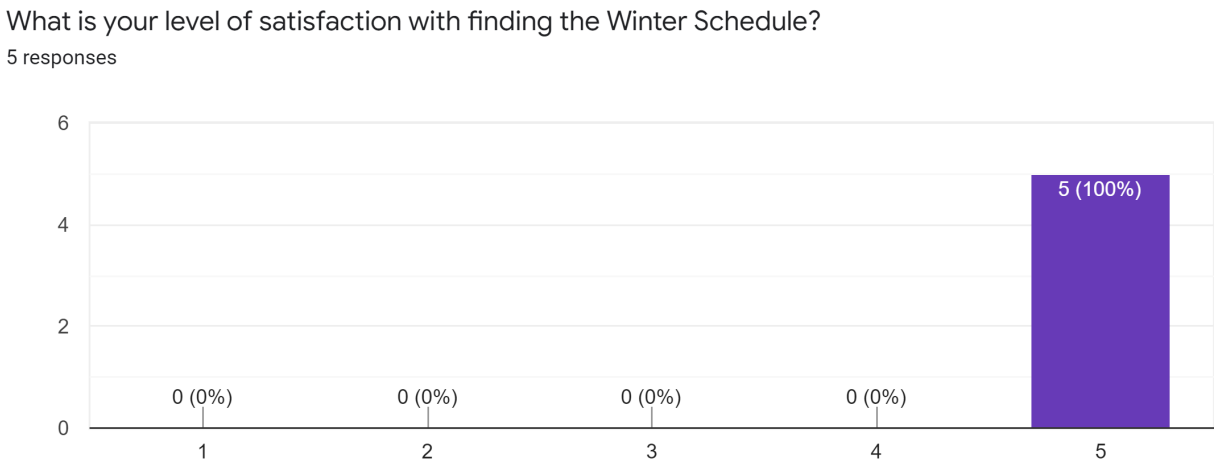
***Task 1C Likert Scale:***




***Task 2 Number of Clicks***

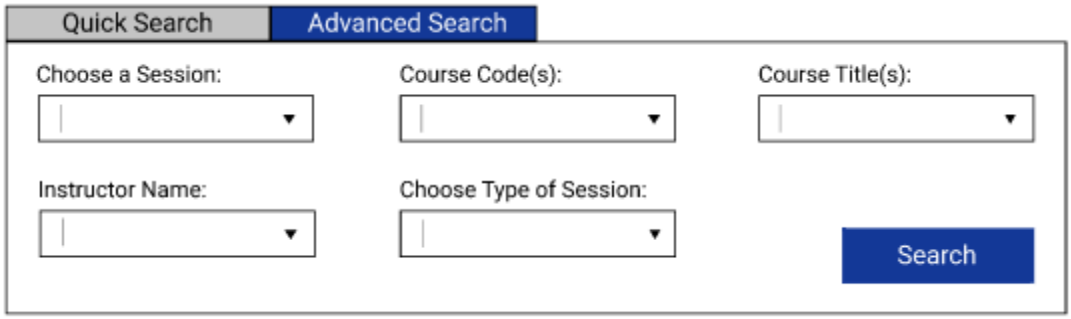
| **Participant interacts with the Timetable as Designed (Accuracy)** | |
|---|---|
| Participant | TASK 2: Number of Clicks to Switch to Winter Timetable |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |

***Task 2 Likert Scale:***



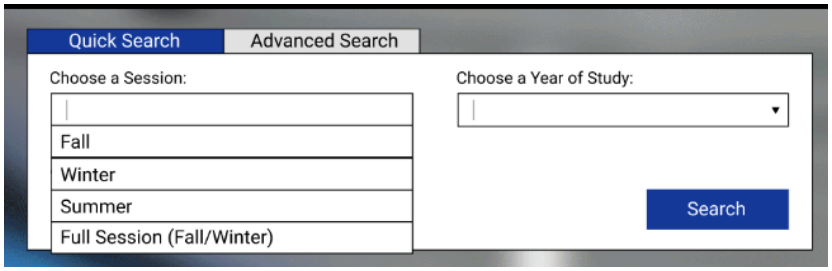

**Changes to Prototype:**

- In our benchmark testing, when asked to add CCT411, users 3 opted to use the advanced search option to look up said course using the course code instead of finding it through the list of 4th year CCT courses provided by the quick search. Therefore we added **Advanced Search** in this iteration of our prototype. (Note: this feature has not been fully implemented.)

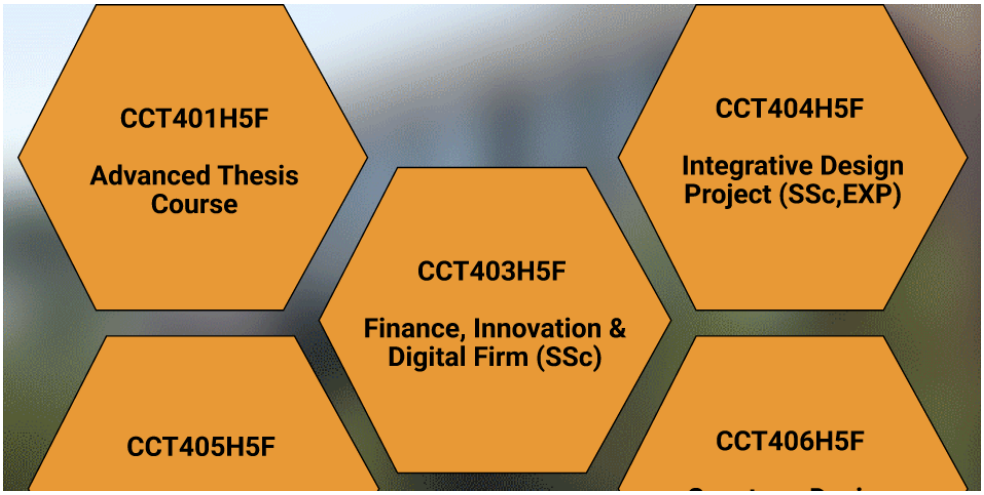



- Originally our 'Choose a Session' section in the Quick Search only offered *Fall*, *Winter*, or *Summer* sessions. In our benchmark testing User 2 and User 3 rated their satisfaction as ⅗ and made comments wondering where the *Full Year* course option was. To solve this

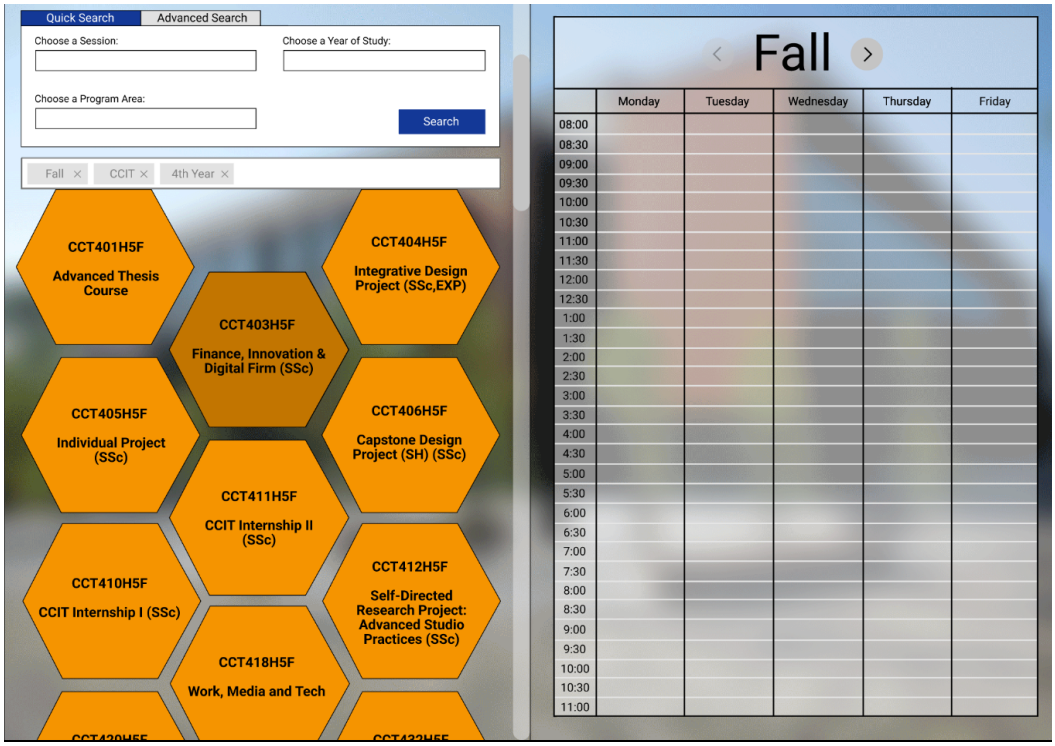
our team **added Y or full session option for session search.**



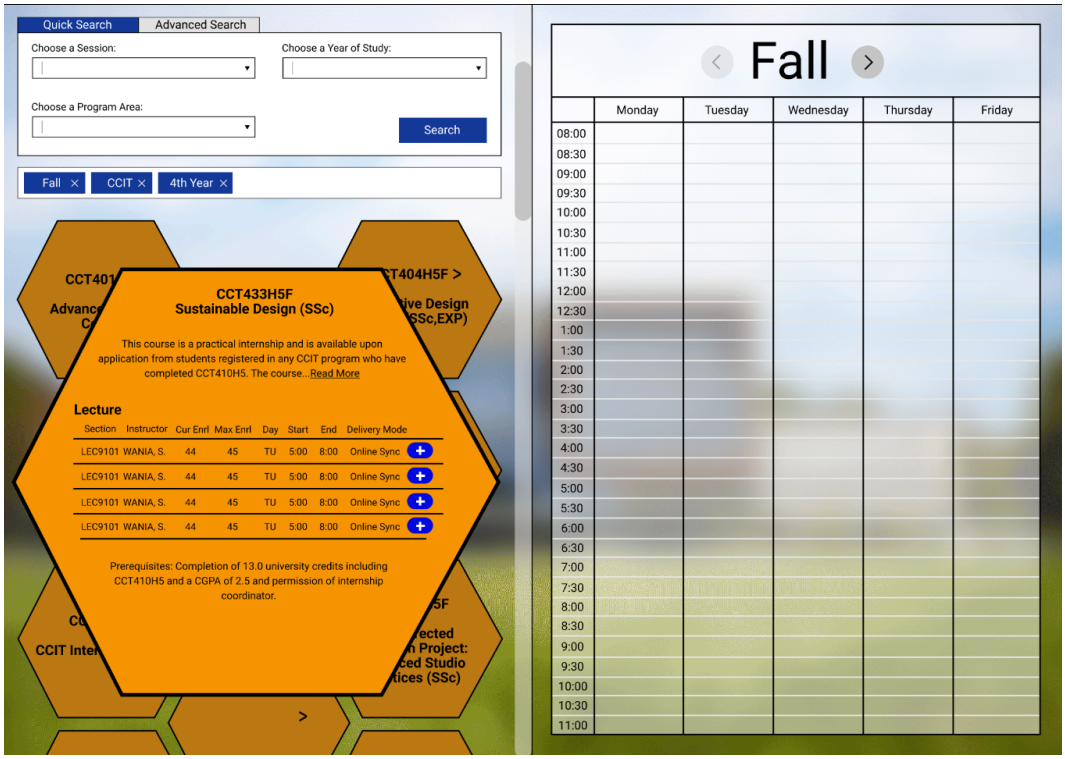

- During feedback from the TA we were made aware that our hexagon courses would not show up for certain screen readers and text would just look like it is not contained as our hexagons did not have outlines, therefore our team **added outlines to our hexagons** for accessibility reading.



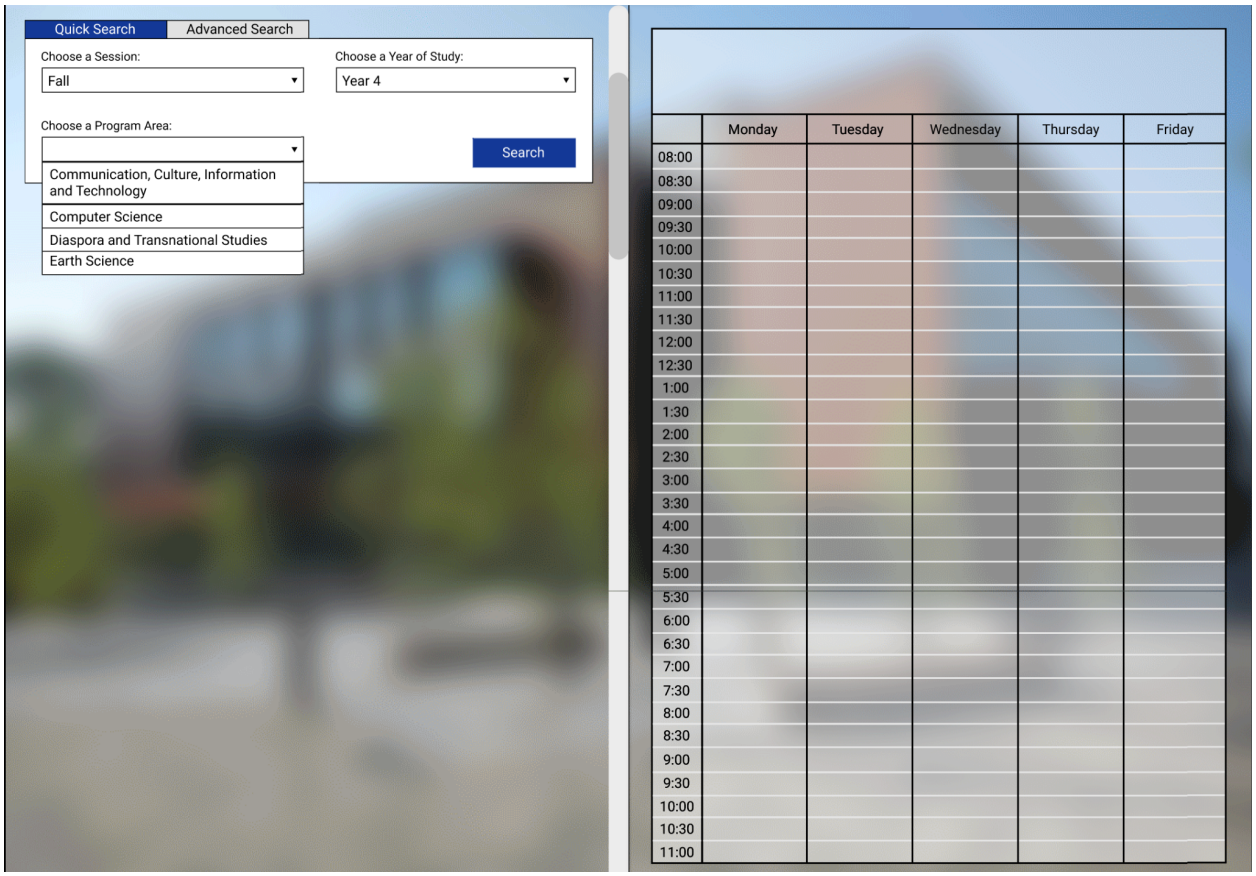

- In our benchmark testing we found all users claimed that they didn't have trouble finding the specific course (CCT411) as it was conveniently centered in placement of the page, giving the likert scale rating of 4 and 5. To create a more dynamic and better understanding of the readability and finding courses we added **hover state to all hexagons.**




- In our benchmark testing we found all users claimed that they didn't have trouble finding the specific course (CCT411) as it was conveniently centered in placement of the page, giving the likert scale rating of 4 and 5. To have better testing we also made **another course have the ability to be added to the schedule** (CCT433).

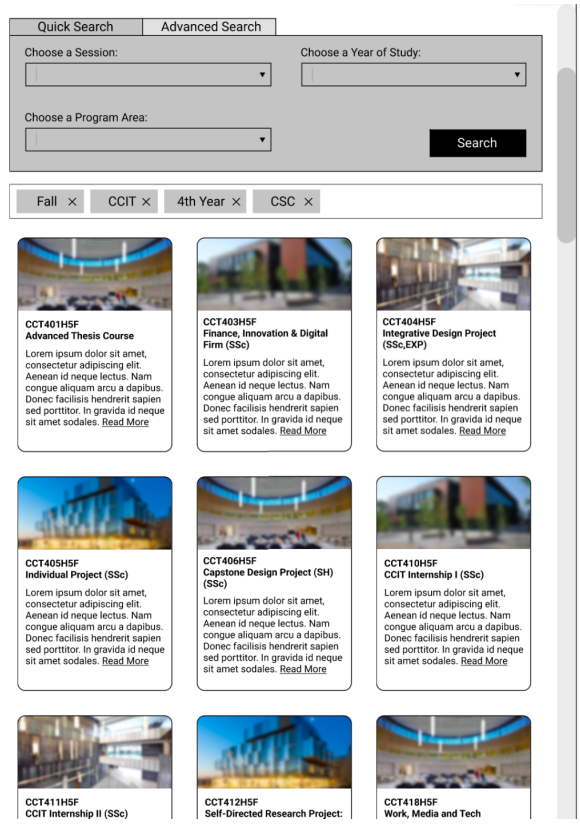


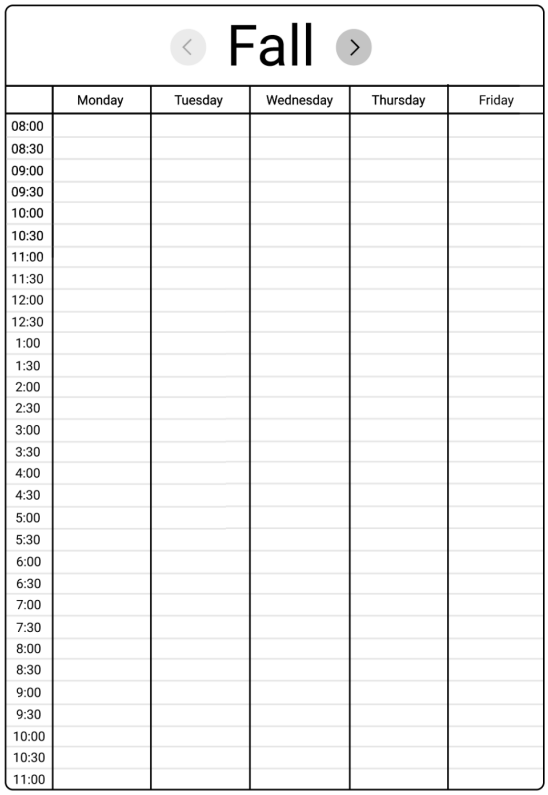
- In our benchmark testing we found all users claimed that they didn't have trouble finding the specific course (CCT411) as it was conveniently centered in placement of the page, giving the likert scale rating of 4 and 5. Therefore, we also added **scrollability** to have more accurate results. Users are now able to scroll down and find more courses.
- In our benchmark testing, we received feedback that the website looked static as if it looked like a diagram rather than something that is interactive. Further, they mentioned that the layout with the timetable filling in half the screen made the website look like two different websites beside each other. To address this, we added a blurred background that transitions through different images of UTM buildings to show more uniformity and movement.

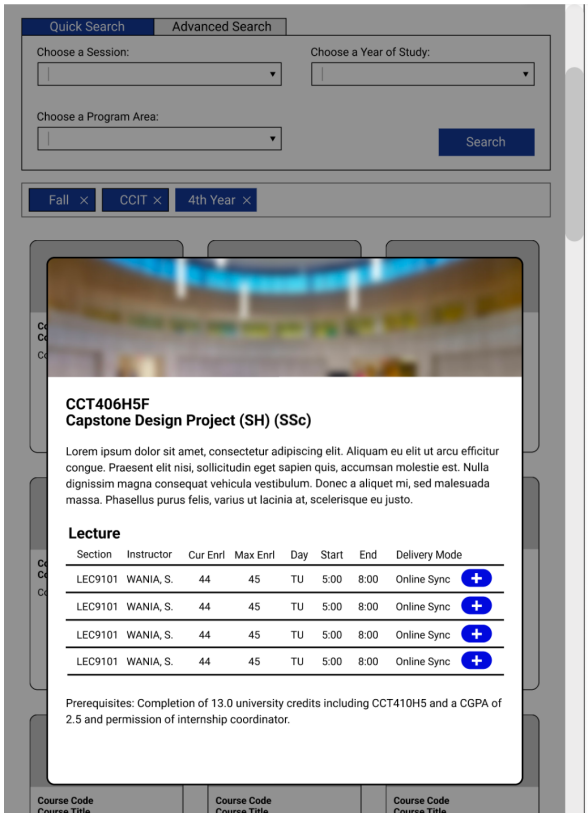


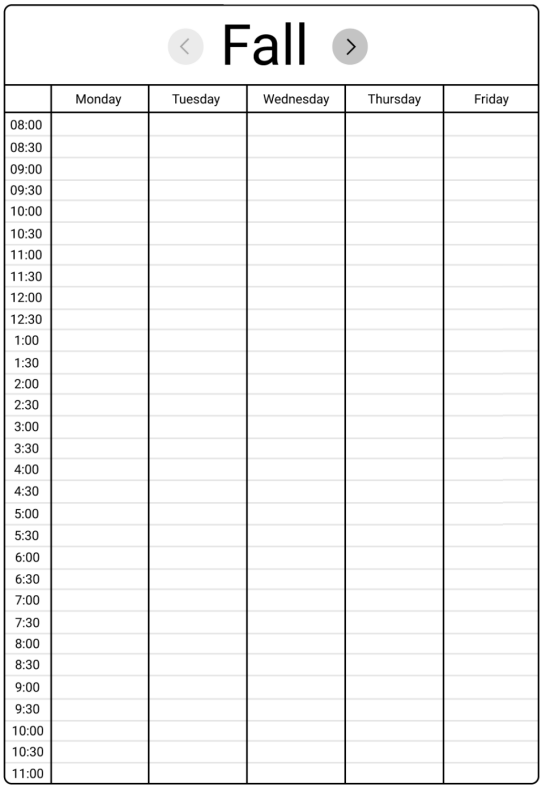

- Changes to benchmark: changed course to add to plan task (to cct433 instead of cct411) our tasks so it requires some scrolling and choose a course not centered to better check the way users scan/read courses.

## Alternative Card Design & Description:



**Description:** The initial design consisted of a search component, a full timetable on the right half of the screen and hexagon panels that represented course options. From our benchmark testing, we found that multiple users struggled in scanning through the courses in a zig-zag pattern that they are not used to. To combat this with the alternative design, we opted with using playing card-like panels for the main interaction of adding a course to the timetable. Each card consists of a demonstrative image of the course concepts, the course title, and the course description.

## Assignment 7: Descriptive Statistics Report on Benchmarking Data

***Note**** *Data Types collected/measured were talked about in Assignment 6 above, Tasks Asked & Data Types Measured section.*

### Likert Scale Data:

| | Participant 1 | Participant 2 | Participant 3 | Participant 4 | Participant 5 |
|---|---|---|---|---|---|
| Finding Course | 4 | 5 | 4 | 5 | 5 |
| Adding Course | 4 | 5 | 5 | 5 | 5 |
| Finding Winter Schedule | 5 | 5 | 5 | 5 | 5 |

### Likert Scale Data Analysis:

| | Mean Score | Median Score | Mode | Variance | Standard Deviation |
|---|---|---|---|---|---|
| Finding Course | 4.6 | 5 | 5 | 0.3 | 0.547722558 |
| Adding Course | 4.8 | 5 | 5 | 0.2 | 0.447213595 |
| Finding Winter Schedule | 5 | 5 | 5 | 0 | 0 |

### Number of Clicks Data:

| Task | Participant 1 | Participant 2 | Participant 3 | Participant 4 | Participant 5 |
|---|---|---|---|---|---|
| Number of clicks to Add to Plan | 1 | 1 | 1 | 1 | 1 |
| Number of clicks to switch to Winter timetable | 1 | 1 | 1 | 1 | 1 |

**Added the number of clicks data as we said we measured the number of clicks data but did not actually include it in our original A7.*

### Number of Clicks Data Analysis:

| Task | Mean Score | Median Score | Mode | Variance | Standard Deviation |
|---|---|---|---|---|---|
| Number of clicks to add to plan | 1 | 1 | 1 | 0 | 0 |
| Number of clicks to switch to timetable plan | 1 | 1 | 1 | 0 | 0 |

**Added the number of clicks data analysis as we said we measured the number of clicks data but did not actually include it in our original A7.*

### Measure of Centrality:

- **Question: What is your level of satisfaction with finding a course?**
    - The median and the mode are the same while the mean is slightly less than the median and mode at 4.6, indicating that it's a slight left-skewed distribution, which are affected by participants 1 and 3 who only had satisfaction ratings of 4 when it came to finding a course.
- **Question:What is your level of satisfaction with adding a course?**

    - The median and the mode are the same while the mean slightly less than both at 4.8, indicating that it's a very slight left-skewed distribution that was affected by participant 1's score.
- **Question: What is your level of Satisfaction with finding the winter schedule?**
    - The mean, the median, and the mode are all the same here, indicating that this is a uniform distribution where all users had the same satisfaction rate with finding the winter schedule.
- **Number of clicks to add to plan:**
    - The mean, the median, and the mode are all the same as every participant used only one click to add a course to plan.
- **Number of clicks to switch to winter timetable:**
    - The mean, the median, and the mode are all the same as every participant used only one click to switch to the winter timetable.

**Added the number of clicks measure of centrality as we did not include it in our original A7.*

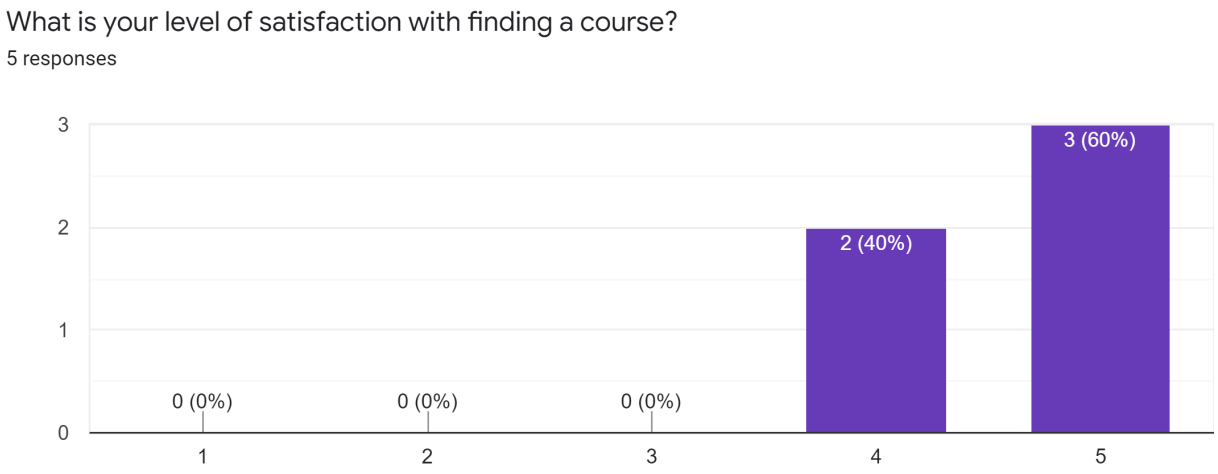
**Measure of Spread:**

- **Question:What is your level of satisfaction with finding a course?**
    - The variance is 0.3, meaning the spread between the numbers is quite low. The standard deviation is 0.5477, indicating that 68% of the data falls between scores of 4.0523 and 5.1477. This is consistent with the data we collected as 3 out of participants rated their experience with this task a 5.
- **Question:What is your level of satisfaction with adding a course?**
    - The variance is 0.2, suggesting that the spread between the numbers is low. The standard deviation sits at 0.4472, meaning that 68% of the data falls between scores of 4.3528 and 5.2472, and 95% of data falls between 3.9056 and 5.6944, which is consistent with the data we collected.
- **Question: What is your level of Satisfaction with finding the winter schedule?**
    - The variance and the standard deviation are both 0, meaning there is no spread and variability with this particular task. This makes sense as all participants gave out 5 as their satisfaction rate.
- **Number of clicks to add to plan:**
    - The variance and the standard deviation are both 0. This is the case as all participants used only one click to add a course to plan and thus our data had no variation.
- **Number of clicks to switch to winter timetable**
    - The variance and the standard deviation are both 0. This is the case as all participants used only one click to switch to the winter timetable and thus our data had no variation.

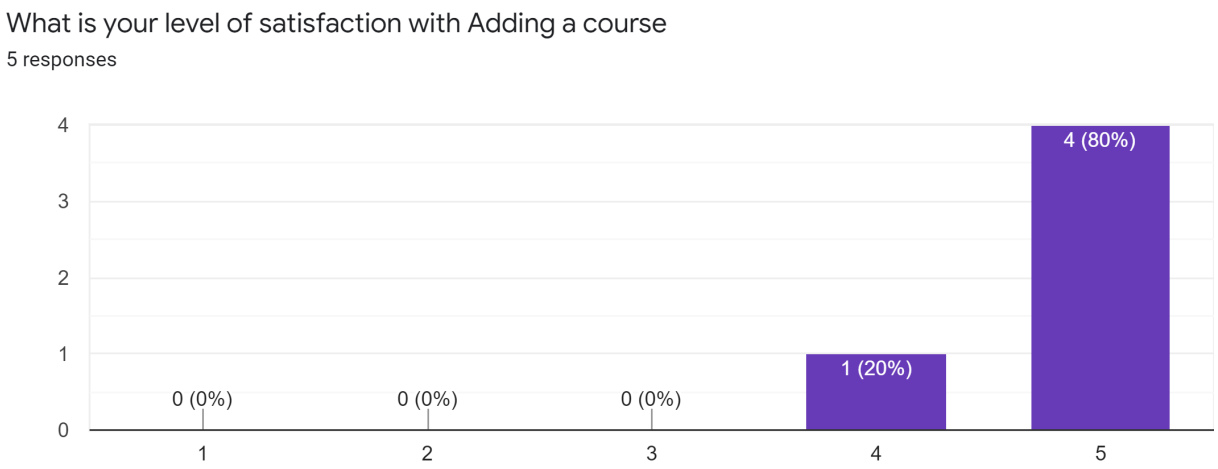
**Added the number of clicks measure of spread as we did not include it in our original A7.*

**Visualizations:**

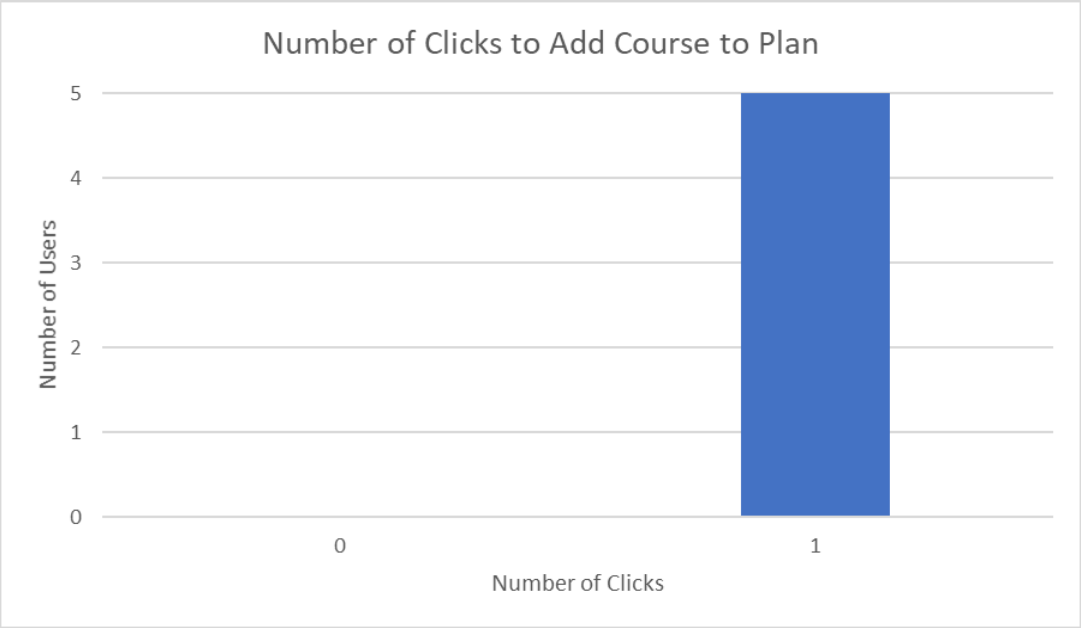
***Task 1B Likert Scale:***




***Task 1C Likert Scale:***

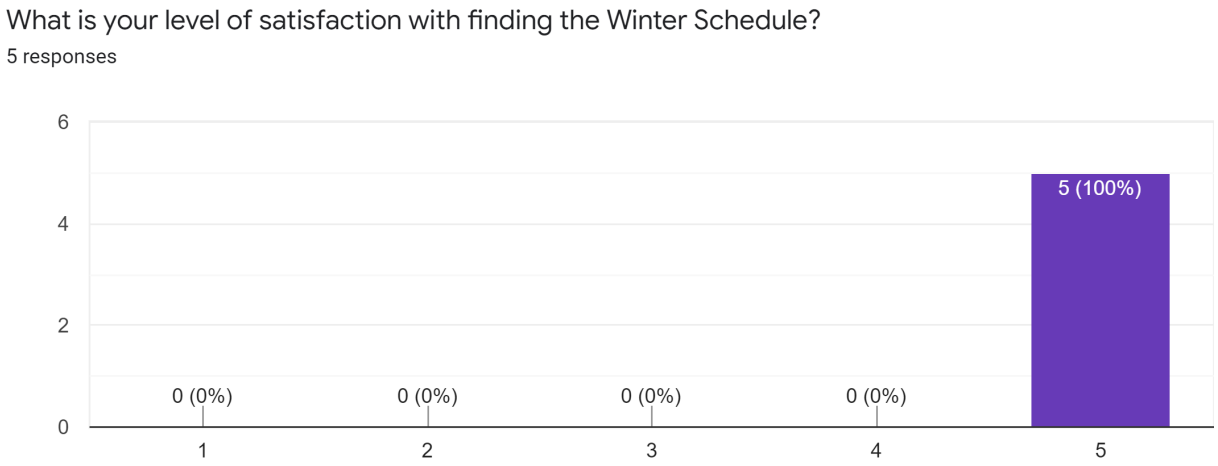


***Task 1C Number of Clicks:***

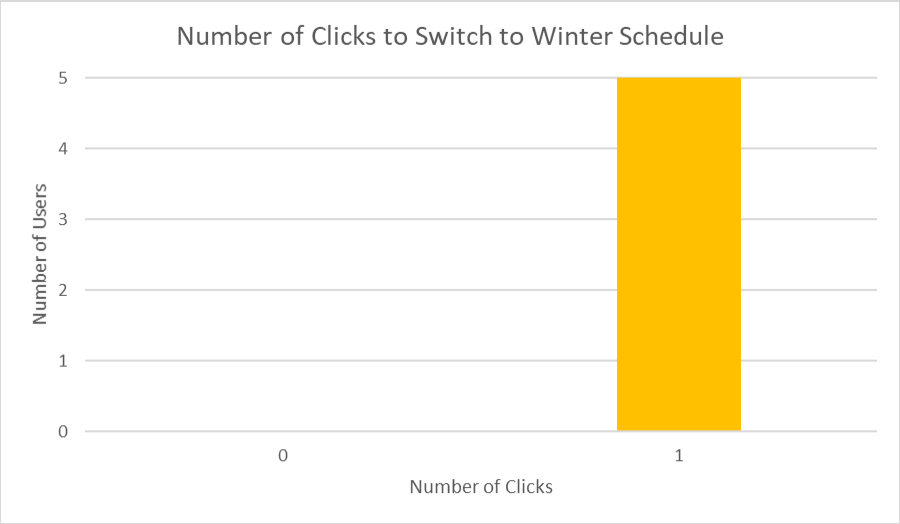



**Added visualizations for number of clicks for task 1c:number of clicks to add course to plan, as we did not include it in our previous submission.*

***Task 2 Likert Scale:***



***Task 2 Number of Clicks:***




**Added visualizations for number of clicks for task 2: switching to winter schedule, as we did not include it in our previous submission.*

**3 Findings Based on Data Analysis**

1. **The final task (checking the winter schedule) has a uniform distribution with all participants rating the task a 5 out of 5. It also has the least variability and deviation from each other and from the mean.**
    a. This makes sense as this task was designed to be straightforward with the goal of testing out the visibility of the arrow design that would allow them to switch between Fall and Winter timetable. The uniformed responses from all 5 participants in this particular sample support the finding that the arrow buttons are visible and easy to use.
2. **As the tasks progress, the mean trends towards the median and the variability and deviation lower as well.**
    a. Although the exact reason for this phenomenon is unclear, it can be due to the fact that the participants became more familiar with the interface as they completed more tasks and were able to figure out where everything was, thus improving their overall experience and satisfaction.
3. **Number of clicks is the same for both adding a course to plan and switching to the winter timetable, with zero variance and standard deviation.**
    a. This makes sense as both tasks should only require 1 click each to complete, meaning none of our users committed any error during these tasks, which is good.

**Added finding for number of clicks for task 1C and task 2 as we changed the quick search task to an onboarding task and therefore needed another finding based on data analysis.*

**A/B Testing Approach:** For A/B testing we will be testing our hexagonal design to our alternative card design. We will be focusing on measuring 'finding the course' part as the difference between our design resides in finding the required course as the hexagonal design and card design follow a different readability design and pattern. We will also be doing counterbalancing the tasks (A: locating course, B: adding course, C: removing course) for each user using both designs.

| **Counterbalancing Latin Squares for Design 1 and 2** | | | |
|---|---|---|---|
| **User** | **Tasks** | | |
| User 1 | A. | B. | C. |
| User 2 | B. | C. | A. |
| User 3 | C. | A. | B. |

Counterbalancing between tasks is necessary as our task A, B, C order is a very expected and predictable process. Counterbalancing will help combat this as it makes the process less predictable for each user. This also helps in avoiding transfer effect given that the order of tasks per user are not necessarily aligned with the prototypes's conventional process such as users expecting a particular task after the other (ie. users would expert the next task would be to select a course after finding it). Also counterbalancing will help us get even results for all tasks incase of limitations in time and users.

**Changed from A7: We changed our approach to include counterbalancing and added the latin squares for task order of design 1 and 2 as feedback from TA. We originally did not do counterbalancing as we were confused about what it was.*

## Assignment 8: Experimental Design & Hypothesis

**Link to New Prototype:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/Untitled?node-id=25%3A139&scaling=scale-down&page-id=0%3A1&starting-point-node-id=25%3A139

**Link To Alternative Design:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/CCT480-UTM-Timetable-Planner?node-id=264%3A157&scaling=scale-down&page-id=5%3A3&starting-point-node-id=264%3A409

**Hypothesis 1:** The A design (alternative card) will have higher accuracy (in terms of number of errors when adding courses to the plan and time taken to locate a course) compared to B design (original hexagonal).
**Changed from A8 version 1: Instead of saying hypothesis 1: one design is more **efficient** for finding a course and hypothesis 2: one design will have less number of errors, we changed it to be **accuracy** by combining the two hypothesis by looking at number of errors/num of clicks (# of errors is # of clicks - 1) and time taken to find courses after feedback from TA.*
**This hypothesis was changed for our alternative hypothesis 1 in A11 to be focused on comprehension of task A(:locating a course) only by looking at time taken and SUS scale for task A.*

**Hypothesis 2:** The A design (alternative card) will have a higher usability and satisfaction rate (in terms of the system usability scale) compared to the B design (original hexagonal) for each task.
**Changed from A8 version 1: We originally measured usability and satisfaction rate but did not have any hypothesis addressing it, therefore we added a hypothesis to only focus on usability and satisfaction rate (SUS and overall satisfaction rate) and specified it was for every task.*
**This hypothesis was changed for our alternative hypothesis 2 in A11 to be focused on satisfaction by looking at overall satisfaction rate (similar to likert scale type questions) and is divided up for satisfaction for each task. Note: SUS scale was changed to be only taken for task A:locating course).*

### Independent & Dependent Variables

Independent variables are as follows:

1. The **prototypes** are the user is using (either the alternative card design or the original hexagonal design).
   **Changed from previous A8: We removed courses we tell users to locate and order of prototypes from independent variables after feedback from TA on how our previous*

*independent variables could just be categorized by what prototypes we are using rather than separating them.*

The correlating dependent variable are as follows:

1. **Time** taken to find the course will be affected by which prototype the user is testing as one design is bound to take less time locating a course because of difference in navigation.
2. **Number of errors/clicks** for adding and removing courses will be affected by which prototype the user is testing as one design is bound to have more number of clicks and errors.
3. **Satisfaction and Usability rate** will be affected by which prototype the user is testing as one design is bound to be favored by users as navigation will be easier in one.
   **Changed from previous A8: After changing independent variables we changed dependent variables to suit the independent variable and removed the repeated dependent of variables such as time taken to find course dependent on the (previous a8) independent variable of which course we ask users to locate.*
   **We also separated time and num of clicks into two separate variables as before it was combined into one.*
   **We also added satisfaction and usability rate as we had not included it before even though it is a correlating dependent variable.*

**Dependent Variables Measurements**

We will measure our data via timing the users per task action. We will also have two or three of us measuring the number of clicks taken to complete each task in the case there is a human error on our part of timing. We will record the number of clicks taken to remove and add courses via a table and have one of our members record the number of clicks. All of this will be done after we have asked the user's consent to record their interactions. We will also be doing overall satisfaction through google surveys in the SUS scale and likert scale format.
**Changed from previous A8: We stated the reason for having two time takers.*
**Added number of clicks measurement and overall satisfaction and usability rating as we did not previously state how we would be taking those measurements.*

**Anticipated Confound Variables**

The confounding variables that we will be anticipating are the user's past experience with the current UTM timetable planner. This can potentially affect the dependent variables such as the time taken to find a course and the amount of errors. This is also true for participants that have been tested before with the original prototype. The last confounding variable to note are the users' familiarity with other online dashboards and websites and how that informs the actions they make during A/B testing.

**Statistical Tests for Each Variable**

We will be conducting **two sample t-tests assuming equal variances** for the **difference in means** of time completion and number of errors/clicks between the 2 prototypes.

## Assignment 9: Experimental Protocol

**A/B Testing Protocol:**

***Introduction:***

Hello and welcome. Thank you for taking the time to test our design. What we'll have you do is go through a bunch of tasks for two different designs of our UTM Timetable Planner website. We will ask you to think aloud to describe what you are thinking throughout the process.

We want to ask for your consent if we can record your interaction with the website, your face or audio will not be recorded.

Before we begin we just want to practice the scrolling for both designs, as it is a bit finicky on Figma. **direct them to scroll between cards and cursor*.*

**Changed from A9 version 1 because we noticed an issue with Design A scrolling on Figma which caused delays in our times taken to locate the course so we incorporated a scrolling test for both designs to eliminate this factor.*

***Counterbalancing Order of Tasks***

| Latin Square: Order of Tasks for Both A & B Design | | | |
|---|---|---|---|
| Users | Tasks | | |
| 1, 4, 7, 10 | A.<br>Find lecture times for CCTXXX. | B.<br>Add CCTXXX to schedule. | C.<br>Remove CCTXXX (course to drop) from schedule. |
| 2, 5, 8, 11 | B. | C. | A. |
| 3, 6, 9, 12 | C. | A | B |

***A Design (card) Scenarios/Tasks & Measurements:***

A. **Task:** You really want to find more info and lecture times for CCT432. Show us where you would find the course info and lecture times for CCT432.

**Measurement:**

| | Time Taken to Locate CCT432 in Design A | | |
|---|---|---|---|
| User | Timer 1 | Timer 2 | Average of Both Timers |

| | | | |
|---|---|---|---|
| 1 | 12.85 | 14.93 | 13.89 |
| 2 | 22.98 | 12.50 | 17.74 |
| 3 | 13.74 | 12.66 | 13.2 |
| 4 | 3.25 | 5.60 | 4.43 |
| 5 | 8.36 | 6.24 | 7.3 |
| 6 | 7.43 | 8.80 | 8.12 |
| 7 | 3.64 | 3.16 | 3.4 |
| 8 | 8.31 | 8.26 | 8.29 |
| 9 | 2.99 | 5.16 | 4.08 |
| 10 | 2.98 | 2.93 | 2.96 |
| 11 | 6.20 | 5.26 | 5.73 |
| 12 | 4.90 | 5.30 | 5.1 |

***Reason for Using 2 Time Takers:** We asked participants to notify us when they had located the course by saying "Done" but as not all participants did this, we had two members of our team record the times and took the average of the two to get the most accurate times for our results.*

B. **Task:** You see that CCT432 seems fun and fits into your schedule so you want to add it to your schedule. Show us how you would add a CCT432 lecture to your schedule.
**Measurement:**

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |

| | |
|---|---|
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

C. **Task:** You find that there is a CCT406 course already added on your schedule. You look into it and find it doesn't seem fun and you don't want to take it anymore. Show us how you would remove the course from your schedule.
**Measurement:**

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

***B Design (hexagonal) Scenarios and Tasks :***

A. **Task:** You really want to find more info and lecture times for CCT433. Show us where you would find the course info and lecture times for CCT433.
**Measurement:**

| | Time Taken to Locate CCT433 in Design B |
|---|---|

| User | Timer 1 | Timer 2 | Average of Both Timers |
|---|---|---|---|
| 2 | 10.77 | 9.50 | 10.14 |
| 3 | 5.09 | 4.39 | 4.74 |
| 4 | 3.97 | 3.09 | 3.53 |
| 5 | 5.16 | 5.43 | 5.30 |
| 6 | 4.90 | 5.15 | 5.03 |
| 7 | 8.53 | 9.18 | 8.86 |
| 8 | 10.51 | 9.94 | 10.23 |
| 9 | 10.20 | 8.75 | 9.48 |
| 10 | 7.15 | 7.60 | 7.38 |
| 11 | 4.04 | 4.77 | 4.41 |
| 12 | 5.37 | 4.79 | 5.08 |

***Reason for Using 2 Time Takers:** We asked participants to notify us when they had located the course by saying "Done" but as not all participants did this, we had two members of our team record the times and took the average of the two to get the most accurate times for our results.*

**System Usability Survey (SUS) Scale Survey**

Now we would like to ask you to do a survey, on system usability survey for the task you just completed of locating a course in design X:
https://docs.google.com/forms/d/e/1FAIpQLScpKlJ_4l7nUYULNfWbR5f34OiFRroNp1eKCmJcepr5NZRo7A/viewform

**Changed from previous A9: We changed the sus scale to only be for locating a course (Task A) since we realized that doing it for each individual task would be very tedious. Especially since our tasks were so short. Task A had a different process between both design A (Card design) and design B (Hex Design) so we would have different usability between users. Task B (adding a course) and Task C (removing a course) were the same process in between Design A and Design B so we did not think we would get a difference in usability between the two. This is why we decided to do a SUS scale for Task A and not Task B and C.*

**Task:** You see that CCT433 seems fun and fits into your schedule so you want to add it to your schedule. Show us how you would add a CCT433 lecture to your schedule.

**Measurement:**

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

B. **Task:** You find that there is a CCT411 course already added on your schedule. You look into it and find it doesn't seem fun and you don't want to take it anymore. Show us how you would remove the course from your schedule.
**Measurement:**

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |

| | |
|---|---|
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

**User Satisfaction Survey**

**Now we would like to ask you to do an overall satisfaction survey for each task in design X https://forms.gle/gjganHNZVenNyC7M7.**

**Changed from previous A9: We added a overall satisfaction survey at the end after the user did all the tasks. We didn't add it in between each task for the same reason since it was tedious for users and they all expressed irritation. Since they found it tedious a lot of the users would just put the same rating without even thinking about it. This would not give us accurate data and would be useless in our findings. This was a limitation on our part and we explained it in the limitation section.*

***Conclusion:***

We wanted to thank you for your time and ask if you have any additional feedback or questions for us.

## Assignment 10: A/B Pilot Testing

**Link to New Prototype:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/Untitled?node-id=25%3A139&scaling=scale-down&page-id=0%3A1&starting-point-node-id=25%3A139

**Link To Alternative Design:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/CCT480-UTM-Timetable-Planner?node-id=264%3A157&scaling=scale-down&page-id=5%3A3&starting-point-node-id=264%3A409

**Pilot Testing Results**

***A Design (card) Scenarios/Tasks & Measurements:***

Task A: Time Taken to Locate CCT432 A/B Testing Results

| Time Taken to Locate CCT432 | | | |
|---|---|---|---|
| User | Timer 1 | Timer 2 | Average of Both Timers |
| 2 | 22.98 | 12.50 | 17.74 |
| 3 | 13.74 | 12.66 | 13.2 |
| 4 | 3.25 | 5.60 | 4.43 |
| 5 | 8.36 | 6.24 | 7.3 |
| 6 | 7.43 | 8.80 | 8.12 |
| 7 | 3.64 | 3.16 | 3.4 |
| 8 | 8.31 | 8.26 | 8.29 |
| 9 | 2.99 | 5.16 | 4.08 |
| 10 | 2.98 | 2.93 | 2.96 |
| 11 | 6.20 | 5.26 | 5.73 |
| 12 | 4.90 | 5.30 | 5.1 |

***Reason for Using 2 Time Takers:** We asked participants to notify us when they had located the course by saying "Done" but as not all participants did this, we had two members of our team record the times and took the average of the two to get the most accurate times for our results.*

Task B: Adding a Course from Schedule A/B Testing Results

| Add a Course from Schedule | |
|---|---|
| User | Number of Clicks |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

Task C: Removing a Course from Schedule A/B Testing Results

| Remove a Course from Schedule | |
|---|---|
| User | Number of Clicks |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |

| | |
|---|---|
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

***B Design (hexagonal) Scenarios and Tasks:***

Task A: Time Taken to Locate CCT433 A/B Testing Results

| Time Taken to Locate CCT433 | | | |
|---|---|---|---|
| User | Timer 1 | Timer 2 | Average of Both Timers |
| 2 | 10.77 | 9.50 | 10.14 |
| 3 | 5.09 | 4.39 | 4.74 |
| 4 | 3.97 | 3.09 | 3.53 |
| 5 | 5.16 | 5.43 | 5.30 |
| 6 | 4.90 | 5.15 | 5.03 |
| 7 | 8.53 | 9.18 | 8.86 |
| 8 | 10.51 | 9.94 | 10.23 |
| 9 | 10.20 | 8.75 | 9.48 |
| 10 | 7.15 | 7.60 | 7.38 |
| 11 | 4.04 | 4.77 | 4.41 |
| 12 | 5.37 | 4.79 | 5.08 |

***Reason for Using 2 Time Takers:** We asked participants to notify us when they had located the course by saying "Done" but as not all participants did this, we had two members of our team record the times and took the average of the two to get the most accurate times for our results.*

Task B: Add a Course from Schedule A/B Testing Results

| Add a Course from Schedule |
|---|

| User | Number of Clicks |
|---|---|
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

Task C: Removing a Course from Schedule A/B Testing Results

| Remove a Course from Schedule | |
|---|---|
| User | Number of Clicks |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |

| | |
|---|---|
| 12 | 1 |

***Additional Feedback Used for Prototype Updates***

- **User 1: (Did not record)**
- **User 2:** User 2 thinks that the pictures of the Card Design does not mean anything although they prefer its layout. He found the Card design smaller in comparison to the hexagon design. They prefer to see the prerequisites and then a one sentence summary of the course on the card itself. They conclude that they actually like the hexagon design better and recommend fixing the ordering of the course.
- **User 3:** (Did not record)
- **User 4:** User 4 enjoyed both of them and thought both designs were simple. Has no recommendations for changes.
- **User 5:** Finds the hexagons confusing and they like the Card Design better.
- **User 6:** Like User 5, User 6 likes the card design more. With the Hexagon Design, they find reading on the hexagons more complicated.

**Changed from original A10: We added the additional feedback users provided us at the end of the A/B testing.*

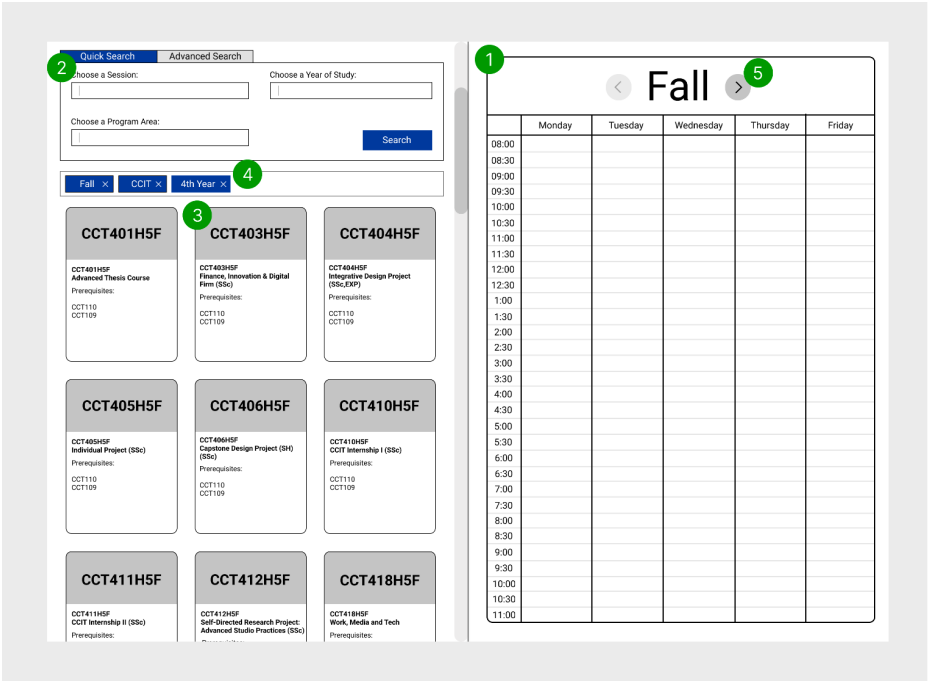
**Changes to Protocol**

Our pilot testing showed us that Users 1, 2, and 3 struggled to scroll on our Design A (card) and in turn their time was greatly affected. Scrolling on Design A is a bit finicky on Figma so this turned out to be a confounding variable and affected our dependent variable of time taken to find the course. So we **added a scrolling onboarding task in our protocol to instruct users on how to properly scroll.** For protocol, we also **only tested system usability scale and overall satisfaction at the end of each design's 3 tasks** instead of once after every test as we found that the survey after each task was tedious for users and they expressed irritation. As users found the survey tedious they would often just put the same rating without thinking about it.

**Changes to Prototypes**

Another thing we fixed was **ordering of courses on Design B (hex) prototype** as there was a human error on our team's part about courses not being in numeric order. This (confounding variable) also affected the time taken to find the course (dependent variable). We also changed our **Design A (card) prototype by replacing the generic UTM scenery images for each course and replacing it with course code.** Another thing we changed on **Design A (card) prototype was changing the course description on each card preview and replacing it with prerequisites.**

## Annotated Prototypes Design A (alternative card design) with Changes

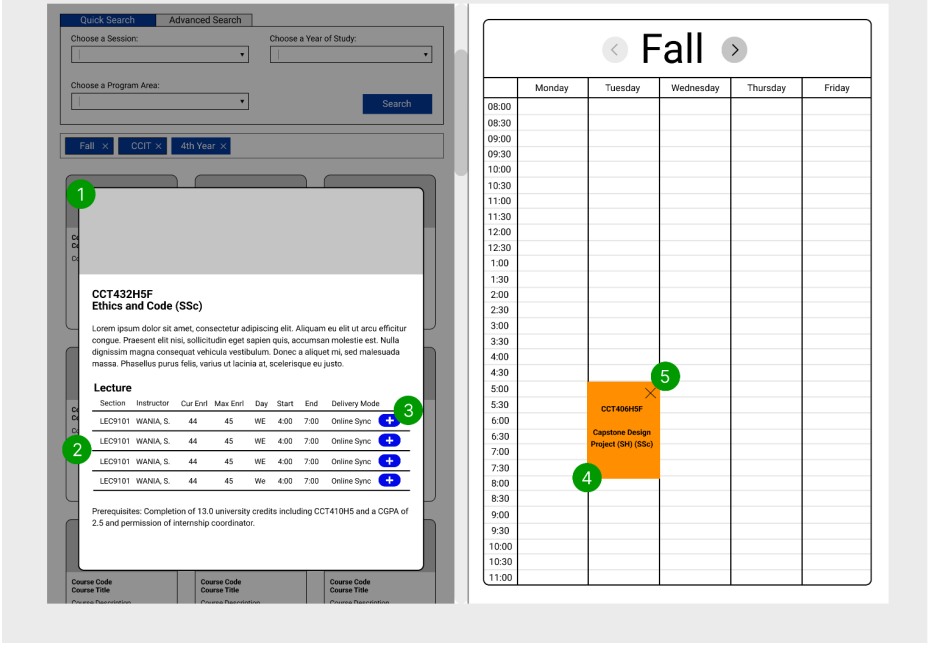



**Design A - Alternative Card Design**

Annotations

1. The calendar for each user in which courses are displayed when the user adds them to their schedule.
2. The Quick Search that the users can use to find courses. They can choose between which session, Year of study, and Program Area. Once they click "Search" the content is loaded.
3. Each individual course shown in a card design. You can see the course code, the course title, and the prerequisites.
4. The filters that the user did to search for courses. they can click the "X" to quickly remove them.
5. The Arrows once clicked changes the calendar to the winter session calendar. Once they are on the winter session calendar, they can click the other arrow to go back.

For our Design A (card) prototype we changed the generic UTM scenery images for each course and replaced it with the course code. We also changed the course description on each card preview and replaced it with prerequisites.




**Design A - Alternative Card Design**

Annotations

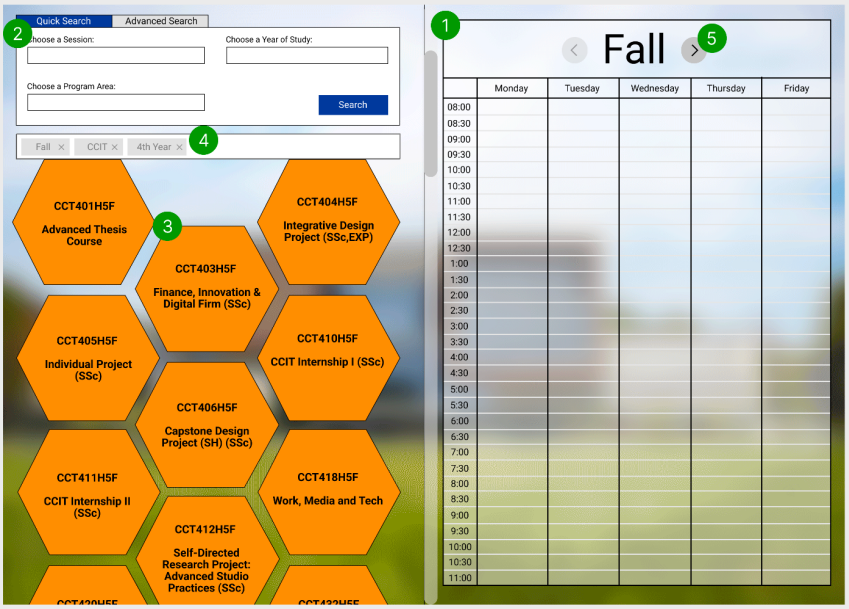
1. Once the user clicks onto a course they are prompted with this card of information. This card includes the Course Code, Course name, Course Description, Lecture information, prerequisites, etc.
2. Lecture infromation that includes different lecture Sections, Instructors, Current Enroled Students #, Max Enrolment #, Start time, End time, and Delivery mode.
3. The user clicks on the "plus" button to add the course (relative to the session) on to the course schedule/calendar.
4. Once the user adds a course, it gets added to the calendar. The course populates the time it is occuring and has the course title and name on it.
5. The user can click the "X" button to quickly remove the course from the course schedule/calendar.

We changed the generic UTM scenery images to a gray placeholder for the lecture times and info popup users see when they click on a course.

**Changed from A10: we added our design A annotated prototypes also stating any changes we made to the design.*

## Annotated Prototypes Design B (original hex design) with Changes

**Changed from A10: we added our design B annotated prototypes also stating any changes we made to the design.*

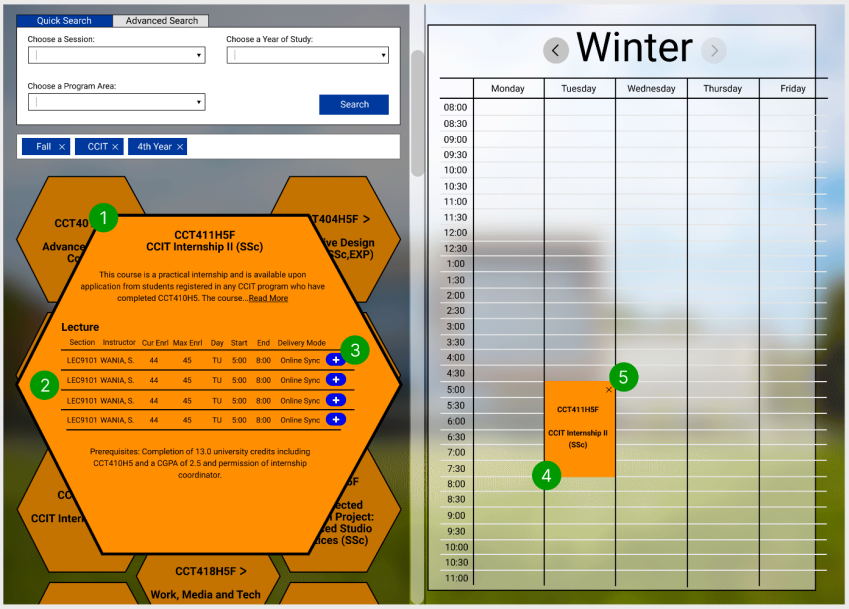



**Design B - Original Hex Design**

Annotations

1. The calendar for each user in which courses are displayed when the user adds them to their schedule.
2. The Quick Search that the users can use to find courses. They can choose between which session, Year of study, and Program Area. Once they click "Search" the content is loaded.
3. Each individual course shown in a Hexagonal design. You can see the course code and the course title.
4. The filters that the user did to search for courses. they can click the "X" to quickly remove them.
5. The Arrows once clicked changes the calendar to the winter session calendar. Once they are on the winter session calendar, they can click the other arrow to go back.

We fixed the ordering of courses on the Design B (hex) prototype due to a human error on our team's part about courses not being in numeric order. This (confounding variable) also affected the time taken to find the course (dependent variable).




**Design A - Alternative Card Design**

Annotations

1. Once the user clicks onto a course they are prompted with this card of information. This card includes the Course Code, Course name, Course Description, Lecture information, prerequisites, etc.
2. Lecture infromation that includes Section, Instructor, Current Enrolment, Max Enrolment, Start time, End time, and Delivery mode.
3. The user clicks on the "plus" button to add the course (relative to the session) on to the course schedule/calendar.
4. Once the user adds a course, it gets added to the calendar. The course populates the time it is occuring and has the course title and name on it.
5. The user can click the "X" button to quickly remove the course from the course schedule/calendar.

## Assignment 11: A/B Testing Results

**Research Question:** Which UTM timetable design out of Design A (alternative card) and Design B (hexagonal design) is better in terms of user's **comprehension** when reading through different courses to locate courses and considered more **usable and satisfying** for users.
**Changed from A11 version 1: originally had the research question focused on the difference on **accuracy** for time and adding course to plan and then for **usability and satisfaction** between the two designs. We changed accuracy to comprehension (time taken to locate course and SUS scale) for task A as we wanted one part to focus on just locating courses and our limitation that we only conducted SUS scale test for the one task ie. task A.*

**Null Hypothesis 1:** There is **no difference** between Design A (alternative card) and Design B (hexagonal Design) in terms of **comprehension.**
****Changed from A11 version 2:** originally had 2 different null hypotheses which were both biased on which design is better in terms of 1)accuracy and 2)satisfaction, then we changed it to combine comprehension and satisfaction (more broad), this version has no bias and we are focusing on task A comprehension and separated both comprehension and usability into two different null hypotheses after feedback from our TA and Professor.*
****Comprehension** entails, understanding and reading through course cards to locate the course ie. time taken to locate course and sus scale rating for locating course task.*

**Null Hypothesis 2:** There is **no difference** between Design A (alternative card) and Design B (hexagonal Design) in terms of **satisfaction rate for each task**.
****Changed from A11 version 2:** We divided up our null hypothesis for comprehension and usability so we can tackle the satisfaction rate for each task separately in our statistical analysis and interpretations as some tasks agree with the null and some disagree.*

**Alternative Hypothesis 1:** The A design (alternative card) will have higher **comprehension** (in terms of time taken to locate a course **and SUS - system usability**) compared to B design (original hexagonal).
****Changed from A11 version 1:** originally had an alt hypothesis focused on accuracy and combined number of clicks with time to locate course. We separated the alt hypothesis to focus on task A:locating a course ie. time taken for task A and SUS scale for task A (after incorporating feedback from TA).*

**Alternative Hypothesis 2:** The Design A (alternative card) will yield a **higher satisfaction rate compared to Design B (original hexagonal).**
****Changed from A11 version 1 & 2:** originally had that Design A will have a higher satisfaction rate compared to Design B and then broke down alternative hypotheses for each task so we can address each one specifically with its corresponding likert scale type overall satisfaction rate but found this would contradict our null hypothesis so reverted back to original alt hypothesis that design A will have higher satisfaction and decided to tackle our different tasks agreeing or disagreeing with hypothesis in our interpretations and stats interpretation (after consulting with our TA).*

**Alpha-level $\alpha$ = 0.05** as we are constructing a 95% confidence interval for our results.

**Independent variables**:

1. The **prototypes** are the user is using (either the alternative card design or the original hexagonal design).

**Dependent Variable:**

1. **Time** taken to find the course will be affected by which prototype the user is testing as one design is bound to take less time locating a course.
2. **Number of errors/clicks** will be affected by which prototype the user is testing as one design is bound to have more number of clicks and errors.
3. **SUS scale** for locating the course will be affected by which prototype the user is testing as one design is bound to have a (average) higher sus rate than the other.
4. **Overall satisfaction rate** will also be affected by which prototype the user is testing as one design is bound to be more favored by the user in terms of design, easability of locating course, and other factors.
   **Changes from A11: Added sus scale overall satisfaction rate as they are also dependent variables but were not stated in previous submission.*

**Confounding Variables:** The confounding variables that we will be anticipating are the **user's past experience** with the current UTM timetable planner. This can potentially affect the dependent variables such as the time taken to find a course and the amount of errors. This is also true for participants that have been tested before with the original prototype. The last confounding variable to note are the users' familiarity with other online dashboards and websites and how that informs the actions they make during A/B testing.

**Type 1 Error**

- **Overall users have a higher overall satisfaction rate for adding the course for Design A, the card design, with a 4.75 average rating compared to Design B, hex design, with a 4.67 average rating.** This rejects the null hypothesis that there is no difference between Design A and Design B in terms of satisfaction for task B (even though there is no difference in task B between both designs).

**Type 2 Error**

- We did not specifically run into this but a possible type 2 error could be**: overall users gave the same SUS score for task A (locating the course) for Design A (card design) and Design B (hexagonal design).** This rejects the null hypothesis that there is no difference between Design A and Design B in terms of comprehension even though we know that Design A and B are different ie. the null is false.

**Changed from A11: We changed our type 1 and type 2 errors to include overall rather than instances with specific users and changed them to suit the new null hypothesis of there not being a difference between the two designs after getting clarification with TA and Prof.*

**Variable Data Type**

- Prototype Design A or B: Nominal Data Types
- Time: Ratio Data Type
- Number of Clicks: Interval Data Type
- Sus Rate: Ordinal Data Type
- Overall Satisfaction Rate: Ordinal Data Type

**Choice Statistical Test**

We chose paired-samples t-test (since the sample size is less than 30) to compare and test the means of the data between two prototypes. This data includes the time taken to locate a course, the number of clicks taken to add a course, number of clicks taken to remove a course and the user's input on our SUS scale with regards to locating a course and our Overall Satisfaction Survey. The result of this test will inform us whether or not the null hypothesis can be rejected.

**Summary Stats & Stats Test for Task A: Locate a Course Time Taken**

***Summary Statistics for Task A: Time Taken to Locate a Course Design A***

| | |
|---|---|
| Mean | 7.85333 |
| Median | 6.515 |
| Mode | n/a |
| Standard Deviation | 4.721 |
| Variance | 22.2878 |

**Interpretation:**

**Mean, median mode:** Since the Mean is higher than the Median, our data could be skewed to the right. Due to a couple of outliers in our data, it could have pulled the mean to the higher side. For example, with our Figma scrolling issue that occurred with our first couple participants. The high locating time due to them not being able to scroll could have made the mean higher. The reason our Mode is N/A, is due to our timings being to the second decimal point. If we rounded up our data, we would probably have an actual Mode rather than N/A.

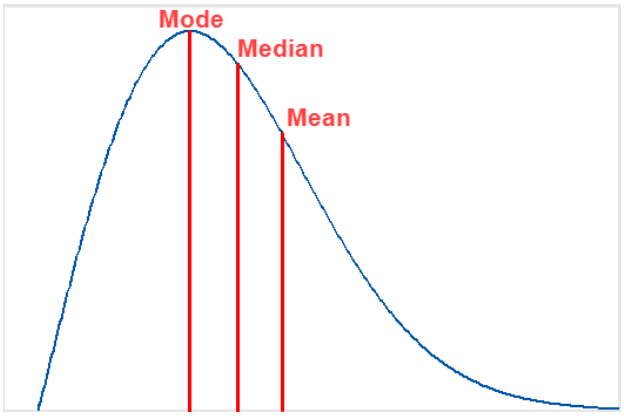



Image received from: https://statisticsbyjim.com/basics/skewed-distribution/

**Standard Deviation & Variance:** Given that the Mean is 7.85 and standard deviation is 4.7 most of the data points should be found in +4 .7 of 7.85 and in -4.7 of 7.85. Thus most data points would be found in the range of 3.15 to 12.55.
**Changed from previous A11: we added summary statistics, stats interpretations, and visualizations as we had not included this in the previous submission.*

***Summary Statistics for Time Taken to Locate a Course Design B***

| | |
|---|---|
| Mean | 7.3275 |
| Median | 6.34 |
| Mode | n/a |
| Standard Deviation | 3.14261 |
| Variance | 9.87602 |

**Interpretations:**
**Mean, median, mode:** Similar to Design A, the Mean is higher than the Median for Design B so our data could be skewed to the right. Due to a couple of outliers in our data, it could have pulled the mean to the higher side. For example, with our Figma scrolling issue that occurred with our first couple participants. The high locating time due to them not being able to scroll could have made the mean higher. The reason our Mode is N/A, is due to our timings being to the second decimal point. If we rounded up our data, we would probably have an actual Mode rather than N/A.

Image received from: https://statisticsbyjim.com/basics/skewed-distribution/

**Standard Deviation & Variance:** Given that the mean is 7.3 and the standard deviation is 3.14 most of the data points should be found in +3.14 of 7.3 and -3.14 of 7.3. Thus most data points would be found in the range of 4.16-10.44.
**Changed from previous A11: we added summary statistics, stats interpretations, and visualizations as we had not included this in the previous submission.*

**Statistical Tests for Time Taken to Locate a Course Design A & B**

***Summary Test for Time Taken to Locate a Course***

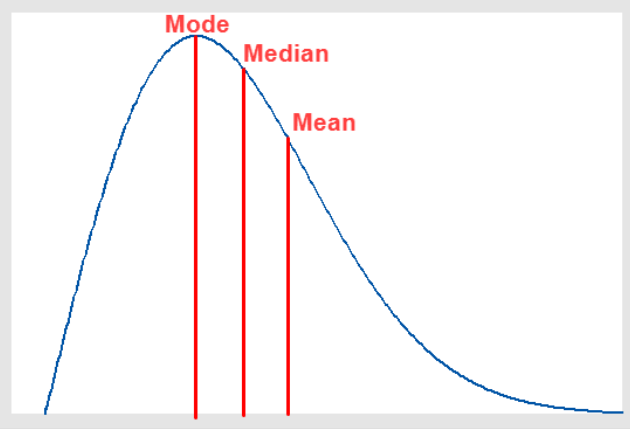



Image received from: https://statisticsbyjim.com/basics/skewed-distribution/

| | *Design A* | *Design B* |
|---|---|---|
| Mean | 7.853333333 | 7.3275 |
| Variance | 22.28780606 | 9.876020455 |
| Observations | 12 | 12 |
| Pooled Variance | 16.08191326 | |
| Hypothesized Mean Difference | 0 | |
| df | 22 | |
| t Stat | 0.321184723 | |
| P(T<=t) one-tail | 0.37555089 | |
| t Critical one-tail | 1.717144374 | |
| P(T<=t) two-tail | 0.75110178 | |
| t Critical two-tail | 2.073873068 | |

**Summary Stats & Stats Test for Task A: Locate a Course SUS Scale**

**Summary Statistics for System Usability Scale - Locating the course - Design A**

| | |
|---|---|
| Mean | 92.9167 |
| Median | 96.25 |
| Mode | 97.5 |
| Standard Deviation | 7.05766 |
| Variance | 49.8106 |

**Interpretation:** Given that the mean is lower than the median, the data will be skewed to the left. Further, the median and the mode are approx one of each other mean but are not too far apart to show much variance.

**Changed from previous A11: we added summary statistics, stats interpretations, and visualizations as we had not included this in the previous submission.*

**Summary Statistics for System Usability Scale - Locating the course - Design B**

| | |
|---|---|
| Mean | 85.4167 |
| Median | 88.75 |
| Mode | 80 |
| Standard Deviation | 13.2216 |
| Variance | 174.811 |

**Interpretation:** Like Design A, the mean is lower than the median meaning that the data can be skewed to the left. In contrast, there is a large gap between the median and the mode and a large variance meaning that the data points are more scattered.

**Changed from previous A11: we added summary statistics, stats interpretations, and visualizations as we had not included this in the previous submission.*

**System Usability Scale Statistical Test**

| | *Card Design SUS* | *Hex Design SUS* |
|---|---|---|
| Mean | 92.91666667 | 85.41666667 |
| Variance | 49.81060606 | 174.8106061 |
| Observations | 12 | 12 |
| Pooled Variance | 112.3106061 | |
| Hypothesized Mean Difference | 0 | |
| df | 22 | |
| t Stat | 1.733510606 | |
| P(T<=t) one-tail | 0.048499863 | |
| t Critical one-tail | 1.717144374 | |
| P(T<=t) two-tail | 0.096999727 | |
| t Critical two-tail | 2.073873068 | |

**Changed from previous A11: we added system usability scale statistical test for locating a course between design A and design B as we had not included this in the previous submission.*

**Summary Stats & Stats Test for Task B: Adding a Course to Plan Number of Clicks**

***Summary Statistics for number of clicks to add a course for Design A***

| | |
|---|---|
| Mean | 1 |
| Median | 1 |
| Mode | 1 |
| Standard Deviation | 0 |
| Variance | 0 |

**Interpretation:** The mean, median, and mode are all 1 since all the users required only one click to add the course. This also explains why the standard deviation and variance are zero as well as there are no variations in our data.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

***Summary Statistics for number of clicks to add a course for Design B***

| | |
|---|---|
| Mean | 1 |
| Median | 1 |
| Mode | 1 |
| Standard Deviation | 0 |
| Variance | 0 |

**Interpretation:** The mean, median, and mode are all 1 since all the users required only one click to add the course. This also explains why the standard deviation and variance are zero as well as there are no variations in our data.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Summary Stats & Stats Test for Task C: Removing a Course to Plan Number of Clicks**

***Summary Statistics for number of clicks to remove a course for Design A***

| | |
|---|---|
| Mean | 1 |
| Median | 1 |
| Mode | 1 |
| Standard Deviation | 0 |
| Variance | 0 |

**Interpretation:** The mean, median, and mode are all 1 since all the users required only one click to remove the course. This also explains why the standard deviation and variance are zero as well as there are no variations in our data.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

***Summary Statistics for number of clicks to remove a course for Design B***

| | |
|---|---|
| Mean | 1 |
| Median | 1 |
| Mode | 1 |
| Standard Deviation | 0 |
| Variance | 0 |

**Interpretation:** The mean, median, and mode are all 1 since all the users required only one click to remove the course. This also explains why the standard deviation and variance are zero as well as there are no variations in our data.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Summary Stats & Stats Test for Overall Satisfaction for Each Task**

**Summary Statistics for Overall Satisfaction - Locating the Course - Design A**

| | |
|---|---|
| Mean | 4.75 |
| Median | 5 |
| Mode | 5 |
| Standard Deviation | 0.45227 |
| Variance | 0.20455 |

**Interpretation:** Given that the median is higher than the mean, the majority of our data points could be skewed to the left. This could be due to some outliers. In contrast, the median and the mode both showed consistency in relaying a score of 5 along with a standard deviation of less than 1.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Summary Statistics for Overall Satisfaction - Locating the Course - Design B**

| | |
|---|---|
| Mean | 3.75 |
| Median | 4 |
| Mode | 4 |
| Standard Deviation | 0.86603 |
| Variance | 0.75 |

**Interpretation:** Like Design A, the median is higher than the mean which can mean that most data points skew to the left. Although, the median and the mode show a consistency of a score of 4 rather than 5.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Overall Satisfaction Scale Statistical Test - Locating a Course**

| | *Satisfaction Rating for Design A Task A* | *Satisfaction Rating for Design B Task A* |
|---|---|---|
| Mean | 4.75 | 3.75 |
| Variance | 0.204545455 | 0.75 |
| Observations | 12 | 12 |
| Pooled Variance | 0.477272727 | |
| Hypothesized Mean Difference | 0 | |
| df | 22 | |
| t Stat | 3.545621042 | |
| P(T<=t) one-tail | 0.000907022 | |
| t Critical one-tail | 1.717144374 | |
| P(T<=t) two-tail | 0.001814044 | |
| t Critical two-tail | 2.073873068 | |

**Changed from previous A11: we added task A overall satisfaction between Design A and Design B statistical T test as we had not included this in the previous submission.*

**Summary Statistics for Overall Satisfaction - Adding the Course - Design A**

| | |
|---|---|
| Mean | 4.66667 |
| Median | 5 |
| Mode | 5 |
| Standard Deviation | 0.49237 |
| Variance | 0.24242 |

**Interpretation:** The mean is lower than the median thus the data points may be skewed to the left. In contrast, the median and the mode show a consistency with a score of 5. Thus, the mean may be influenced by some outliers.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Summary Statistics for Overall Satisfaction - Adding the Course - Design B**

| | |
|---|---|
| Mean | 4.75 |
| Median | 5 |
| Mode | 5 |
| Standard Deviation | 0.45227 |
| Variance | 0.20455 |

**Interpretation:** Design B is in a similar situation with Design A having a mean lower than the median and with the median and mode being the same with each other.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Overall Satisfaction Scale Statistical Test - Adding the Course**

| | *Satisfaction Rating for Design A Task B* | *Satisfaction Rating for Design B Task B* |
|---|---|---|
| Mean | 4.666666667 | 4.75 |
| Variance | 0.242424242 | 0.204545455 |
| Observations | 12 | 12 |
| Pooled Variance | 0.223484848 | |
| Hypothesized Mean Difference | 0 | |
| df | 22 | |
| t Stat | -0.43178777 | |
| P(T<=t) one-tail | 0.335048944 | |
| t Critical one-tail | 1.717144374 | |
| P(T<=t) two-tail | 0.670097888 | |
| t Critical two-tail | 2.073873068 | |

**Changed from previous A11: we added task B overall satisfaction between Design A and Design B statistical T test as we had not included this in the previous submission.*

**Summary Statistics for Overall Satisfaction - Removing the Course - Design A**

| | |
|---|---|
| Mean | 4.83333 |
| Median | 5 |
| Mode | 5 |
| Standard Deviation | 0.38925 |
| Variance | 0.15152 |

**Interpretation:** The mean is lower than the median which means that the data can be skewed to the left. The median and the mode are the same and the variance is very low means there is a high consistency in the overall satisfaction score.

**Changed from previous A11: we added summary statistics and stats interpretations as we had not included this in the previous submission.*

**Summary Statistics for Overall Satisfaction - Removing the Course - Design B**

| | |
|---|---|
| Mean | 4.75 |
| Median | 5 |
| Mode | 5 |
| Standard Deviation | 0.45227 |
| Variance | 0.20455 |

**Interpretation:** Like Design A, the mean is lower than the median which means that the data can be skewed to the left. Similarly, the median and the mode are the same while there is a low variance which means that the score across users are quite consistent.

**Overall Satisfaction Scale Statistical Test - Removing the Course**

| | *Satisfaction Rating for Design A Task C* | *Satisfaction Rating for Design B Task C* |
|---|---|---|
| Mean | 4.833333333 | 4.75 |
| Variance | 0.151515152 | 0.204545455 |
| Observations | 12 | 12 |
| Pooled Variance | 0.178030303 | |
| Hypothesized Mean Difference | 0 | |
| df | 22 | |
| t Stat | 0.483779447 | |
| P(T<=t) one-tail | 0.316661483 | |
| t Critical one-tail | 1.717144374 | |
| P(T<=t) two-tail | 0.633322966 | |
| t Critical two-tail | 2.073873068 | |

**Changed from previous A11: we added task C overall satisfaction between Design A and Design B statistical T test as we had not included this in the previous submission.*

**Report Results**

**T-test result on time taken to locate a course**

- Test statistic: 0.3212
- Critical T value (two-tailed): 2.0739
- P-value: 0.7511

**T-test result on number of clicks to add a course**

- Test statistic: 65535
- Critical T value (two-tailed): 2.0739

**Sus-Scale**

- Average SUS score for design A (card): 92.92
- Average SUS score for design B (hex): 85.42
- Test statistic: 1.7335
- Critical T value (two-tailed): 2.0739
- P-value: 0.0970

**Changed from previous A11: we added the statistical results for system usability scale as we had not included this in the previous submission.*

**Overall Satisfaction Scale**

- Overall Average Satisfaction for Hex Design: 4.41
- Overall Average Satisfaction for Card Design: 4.75
- Average satisfaction for card design task A: 4.75
- Average satisfaction for hex design task A: 3.75
- Test statistic: 3.5456
- Critical T value (two-tailed): 2.0739
- P-value: 0.0018
- Average satisfaction for card design task B: 4.67
- Average satisfaction for hex design task B: 4.75

- Test statistic: -0.4318
- Critical T value (two-tailed): 2.0739
- P-value: 0.6701
- Average satisfaction for card design task C: 4.83
- Average satisfaction for hex design task C: 4.75
- Test statistic: 0.4838
- Critical T value (two-tailed): 2.0739
- P-value: 0.6333

**Changed from previous A11: we added the statistical results for satisfaction scale per task as we had not included this in the previous submission.*

**Results Interpretations**

When looking at the time taken to locate a course, the test statistic (0.3212) was less than the critical value for a two-tailed t-test (2.0739, $\alpha = 0.05$) and the p-value (0.7511) was greater than 0.05, meaning the difference between the means (Design A: card design Mean= 7.853; Design B: hex design Mean = 7.328) of time taken to locate the course was not significant enough, therefore, we failed to reject null hypothesis 1 and could not accept alternative hypothesis 1. This is probably due to the presence of some outliers in our data because of a scrolling issue on Design A's Figma during the pilot A/B testing. For this reason, we modified our protocol to ask users to practice scrolling beforehand to make sure this problem did not occur during the second round of testing but this already produced some substantial outliers on our Design A: card design task A ( time taken to locate course) measurements.
**Changed from previous A11: added outliers explanation to statistical test results interpretation for time taken to locate a course as we*

Similarly, the test statistic for user SUS score for locating a course (1.7335) was less than the critical value for a two-tailed t-test (2.0739, $\alpha = 0.05$) and the p-value (0.0970) was also greater than 0.05, meaning the difference between the two designs (Mean sus score for A/card = 92.92; Mean sus score for B/hex = 85.42) was not statistically significant and thus we failed to reject null hypothesis 1 and could not accept alternative hypothesis 1.
**Changed from previous A11: added interpretation for SUS score statistical test results as we had not included this in the previous submission.*

The test statistic for task A (time taken to locate a course) satisfaction (3.5456) was greater than the critical value for a two-tailed t-test (2.0739, $\alpha = 0.05$) and the p-value (0.0018) was also less than 0.05, allowing us to reject null hypothesis 2 and accept alternative hypothesis 2. This suggests that the difference between mean satisfaction of design A: card design (Mean = 4.75) and design B: hex design (Mean = 3.75) for task A was statistically significant and that design A: card design yielded a higher satisfaction rate than design B: hex design in terms of locating the course.

**Changed from previous A11: added interpretation for task A satisfaction statistical test results as we had not included this in the previous submission.*

The test statistic for task B (adding a course) satisfaction (-0.4318) was not higher than the critical value for a two-tailed t-test (2.0739, **α** = 0.05) and the p-value (0.6701) was much greater than 0.05, meaning the difference between mean satisfactions of design A: card design (4.67) and design B: hex design (4.75) was not statistically significant, therefore we failed to reject null hypothesis 2.
**Changed from previous A11: added interpretation for task B satisfaction statistical test results as we had not included this in the previous submission.*

The test statistic for task C (removing a course) satisfaction (0.4838) was not higher than the critical value for a two-tailed t-test (2.0739, **α** = 0.05) and the p-value (0.6333) was much greater than 0.05, meaning the difference between mean satisfactions of design A: card design (4.83) and design B: hex design (4.75) was not statistically significant, therefore we failed to reject null hypothesis 2.
**Changed from previous A11: added interpretation for task C satisfaction statistical test results as we had not included this in the previous submission.*

**Limitations & Next Steps**
**No Statistical Tests for Number of Clicks for Tasks B and C (Adding & Removing Course):** All users took only one click to complete tasks B and C. In effect, we decided to not include the number of clicks in our null or alternative hypothesis and not conduct statistical tests for these tasks given that the mean, median, and mode would be one and there would be no variance. However, we did include in the summary statistics for consistency sake after consulting without TA.

**No Counter Balancing Between Design A and B:** Our protocol only consisted of counterbalancing between tasks for each design. Ie. User 2 will do tasks in order of task B, task C and task for Design A first and then the same order for Design B. We always conducted Design A tests first. This created a learning transfer effect for users doing Design B as they already know what the tasks entail and can prepare for them. For example users will already start looking at the courses for Design B because they know we will ask them to locate a specific course. This potentially had an effect on our times for task B locating a course. If we were given a chance to redo the tests we would counterbalance between Designs as well as tasks.

**Not Conducting SUS scale for Each Task:** Upon observations during the first A/B testing session, we found that other users found the excessive amount of system usability scale surveys hideous and fatiguing. In order to limit the potential skewness in our data, we decided to only do SUS surveys for locating the course. That being said, system usability scale offers a much more

comprehensive insight than satisfaction likert scales. If given the chance, we would focus entirely on system usability by asking users to fill out SUS surveys after each task and omitting overall satisfaction likert scale completely to limit the overall number of surveys users need to take and reduce the potential fatigue participants may feel.

**Same Process for Task B and C (Adding & Removing Course) in Design A and B:** Another limitation in our project was that task B: removing course and task C:adding course consist of the same steps and processes for both design A and B. This meant that there was no variation in the results between the two; All users took 1 click to add and 1 click to remove the course. This did not help to disprove our null hypothesis at all as both averages were the same ie. 1 click. Our team recognized these two tasks should have been tested in our guerilla testing phase for design B (hex design) rather than A/B testing which would test the difference and compare between two different designs. A/B testing should have been more focused on readability/comprehension/navigation of locating courses (ie. task A:locating course).
**Changed from previous A11: added an extra section for limitations on our A/B testing, statistical tests and the project overall as well as how we would do things differently given the chance.*

## Appendix A: Links to Prototypes

***Link to Wireframes:***
https://manahilmasroor718356.invisionapp.com/freehand/Untitled-Qr5qUjyOT?dsid_h=bc77f8d5d0c43f1e87aef00217919c44bedc7a1e080a1ad33e01d1e3bd76e314&uid_h=accb26bbc3af0748df4a1546bdc88c3fb19b93bc99bc8156d935a0696a6d3e83

**Link to New Prototype:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/Untitled?node-id=25%3A139&scaling=scale-down&page-id=0%3A1&starting-point-node-id=25%3A139

**Link To Alternative Design:**
https://www.figma.com/proto/KWWQtSZcVBVLY7OBqv2Ue7/CCT480-UTM-Timetable-Planner?node-id=264%3A157&scaling=scale-down&page-id=5%3A3&starting-point-node-id=264%3A409

## Appendix B: Raw Data

**Table 1**

***Results for Time Taken to Locate a Course Design A***

| | **Time Taken to Locate CCT432H5 Course in seconds** | | |
|---|---|---|---|
| **User** | **Timer 1** | **Timer 2** | **Average of both Timers** |
| 1 | 12.85 | 14.93 | 13.89 |
| 2 | 22.98 | 12.50 | 17.74 |
| 3 | 13.74 | 12.66 | 13.2 |
| 4 | 03.25 | 5.60 | 4.43 |
| 5 | 08.36 | 6.24 | 7.3 |
| 6 | 07.43 | 8.80 | 8.12 |
| 7 | 03.64 | 3.16 | 3.4 |
| 8 | 8.31 | 8.26 | 8.29 |
| 9 | 02.99 | 5.16 | 4.08 |
| 10 | 02.98 | 2.93 | 2.96 |
| 11 | 06.20 | 5.26 | 5.73 |
| 12 | 04.90 | 5.30 | 5.1 |

**Table 2**

***Results for Time Taken to Locate a Course Design B***

| Time Taken to Locate CCT433H5 Course in Seconds | | | |
|---|---|---|---|
| **User** | **Timer 1** | **Timer 2** | **Average of both timers** |
| 1 | 13.55 | 13.95 | 13.75 |
| 2 | 10.77 | 9.50 | 10.14 |
| 3 | 5.09 | 4.39 | 4.74 |
| 4 | 3.97 | 3.09 | 3.53 |
| 5 | 5.16 | 5.43 | 5.30 |
| 6 | 4.90 | 5.15 | 5.03 |
| 7 | 8.53 | 9.18 | 8.86 |
| 8 | 10.51 | 9.94 | 10.23 |
| 9 | 10.20 | 8.75 | 9.48 |
| 10 | 07.15 | 7.60 | 7.38 |
| 11 | 4.04 | 04.77 | 4.41 |
| 12 | 5.37 | 4.79 | 5.08 |

**Table 3**

***Results for number clicks to add a course for Design A***

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

**Table 4**

***Results for number of clicks to remove a course for Design A***

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

**Table 5**

***Results for number of clicks to add a course for Design B***

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

**Table 6**

***Results for number of clicks to remove a course for Design B***

| User | Number of Clicks |
|---|---|
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 1 |
| 5 | 1 |
| 6 | 1 |
| 7 | 1 |
| 8 | 1 |
| 9 | 1 |
| 10 | 1 |
| 11 | 1 |
| 12 | 1 |

**Table 7**

***System Usability Scale Score - Locating the Course***

| User | Design A (Card Design) | Design B (Hex Design) |
|---|---|---|
| 1 | 97.5 | 90 |
| 2 | 80 | 80 |
| 3 | 100 | 77.5 |
| 4 | 87.5 | 95 |
| 5 | 100 | 65 |
| 6 | 85 | 60 |
| 7 | 90 | 80 |
| 8 | 100 | 97.5 |
| 9 | 97.5 | 100 |
| 10 | 97.5 | 92.5 |
| 11 | 95 | 100 |
| 12 | 85 | 87.5 |

**Table 8**

***Overall Satisfaction Scale - Locating the Course***

| **User** | **Design B (Hex Design)** | **Design A (Card Design)** |
|---|---|---|
| 1 | 4 | 5 |
| 2 | 4 | 4 |
| 3 | 3 | 5 |
| 4 | 3 | 5 |
| 5 | 3 | 5 |
| 6 | 2 | 5 |
| 7 | 4 | 5 |
| 8 | 4 | 5 |
| 9 | 5 | 5 |
| 10 | 4 | 5 |
| 11 | 5 | 4 |
| 12 | 4 | 4 |
| **Average** | **3.75** | **4.75** |

**Table 9**

***Overall Satisfaction Scale - Adding the Course***

| **User** | **Hex Design** | **Card Design** |
|---|---|---|
| 1 | 5 | 5 |
| 2 | 5 | 5 |
| 3 | 5 | 5 |
| 4 | 5 | 5 |
| 5 | 5 | 5 |
| 6 | 4 | 4 |
| 7 | 4 | 4 |
| 8 | 5 | 5 |
| 9 | 5 | 4 |
| 10 | 5 | 5 |
| 11 | 5 | 5 |
| 12 | 4 | 4 |
| **Average** | **4.75** | **4.67** |

**Table 10**

***Overall Satisfaction Scale - Removing the Course***

| User | Hex Design | Card Design |
|---|---|---|
| 1 | 5 | 5 |
| 2 | 5 | 5 |
| 3 | 5 | 5 |
| 4 | 5 | 5 |
| 5 | 5 | 5 |
| 6 | 5 | 5 |
| 7 | 4 | 4 |
| 8 | 4 | 5 |
| 9 | 5 | 5 |
| 10 | 5 | 5 |
| 11 | 5 | 5 |
| 12 | 4 | 4 |
| Average | 4.75 | 4.83 |