

Project CDG

#17 Team CDG

TABLE OF CONTENTS

01

Project
Overview

02

Key
Insights

03

Solution
Overview

04

Technical
Overview

05

Gap
Analysis

06

Moving
Forward

07

Conclusion



01

Project Overview

Overview of team members and project analyses

Our Team Members



SHERMIN

Project Manager



IRIS

Data Scientist



IVY

Data Analyst



HUI XING

Visualization Specialist



SONIA

Data Analyst



SHAZA

Data Scientist

Our Stakeholders



GARY HOW

Point Of Contact for CDG
CDG Data Team Lead

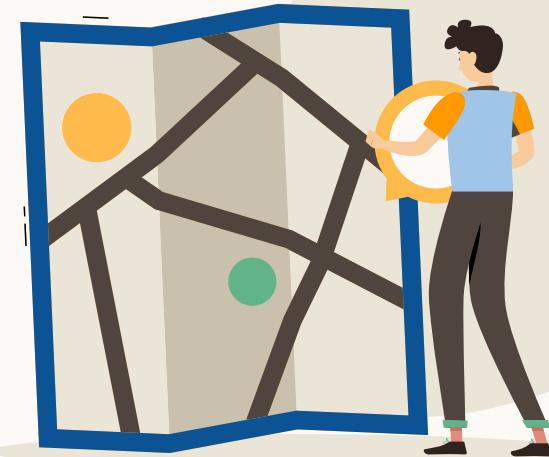


TAN POH CHOO

IS483 Project
Supervisor

PROBLEM STATEMENT

“Discovering the gaps and insights from the demand-influencing factors of the overall Singapore mobility market.”



TARGET AUDIENCE

Tertiary Students in Singapore

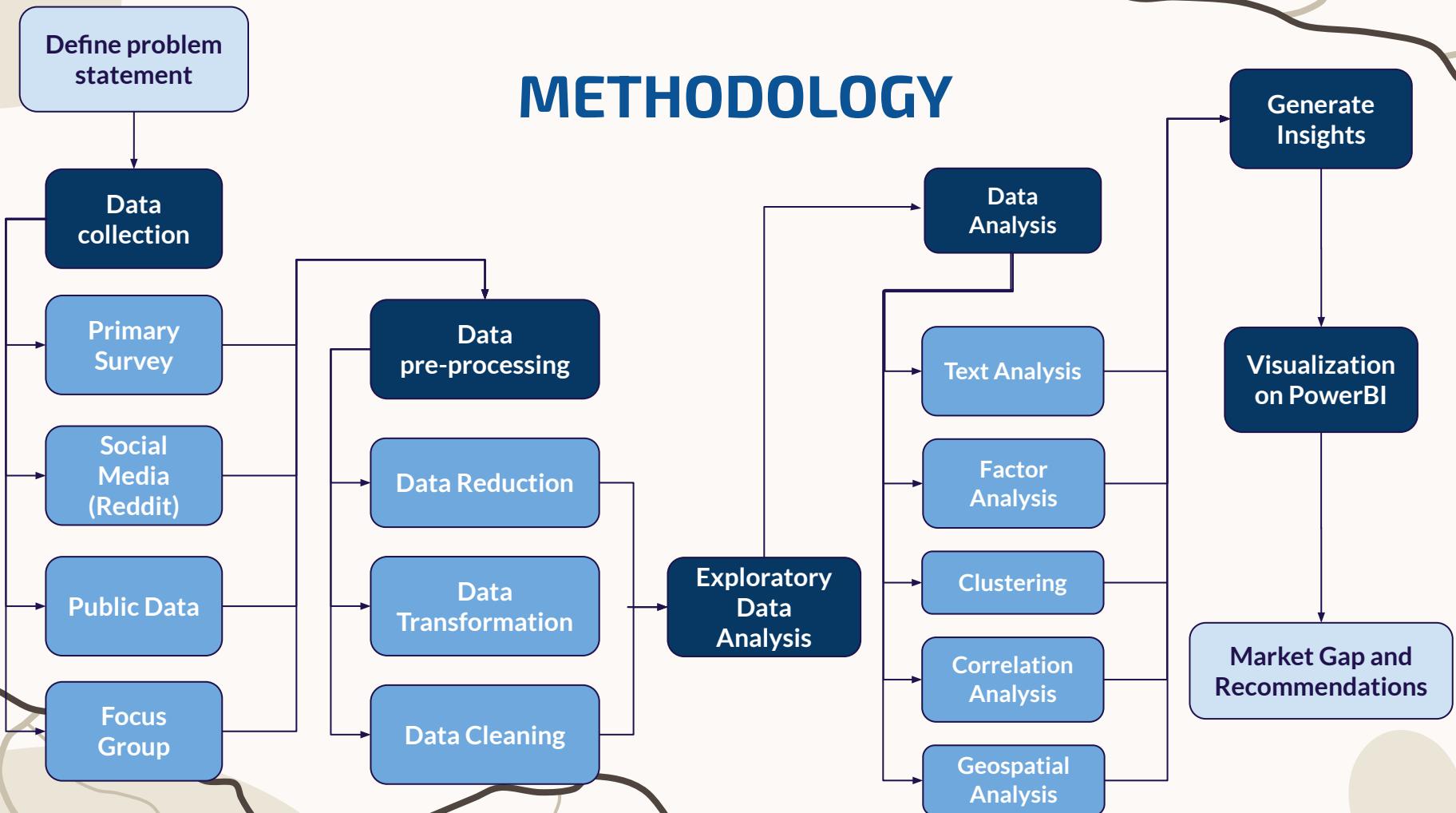


MOTIVATION

Help COMFORTDELGRO gain a better understanding of the Singapore transportation landscape and discover gaps in the current market.



METHODOLOGY



TYPES OF ANALYSIS



Text Analysis



Factor Analysis



Correlation Analysis



Clustering Analysis



Geospatial Analysis

TYPES OF ANALYSIS



Text Analysis

Purpose

Identify the ground sentiments and uncover topical patterns.

How it links to demand

Find out the motivations and disparities of the market.



Factor Analysis

Purpose

Understand the extent to which each variable is associated with a common factor and uncover cluster of responses.

How it links to demand

Identify demand-influencing factors that arise from the survey.



Correlation Analysis

Purpose

Identify the underlying relationship between factors and Customer Satisfaction Index of Singapore (CSISG) scoring.

How it links to demand

Find out main factors that are affecting its demand by sector.

TYPES OF ANALYSIS



Clustering Analysis

Purpose

Identify present clusters amongst the population by demographics and perceptions.

How it links to demand

By visualizing differences from the results, discover gaps in the market.



Geospatial Analysis

Purpose

Identify phenomena or gaps due to changing spatial conditions or location-based events.

How it links to demand

Understand how influencing factors in each planning area are affecting demand.



Key Insights

TEXT ANALYSIS

Key Insights



Most common keywords

Price (money, pay, fare, cost),
time and service

Prominent Topics observed

Ease of travelling and time taken while
travelling

FACTOR ANALYSIS

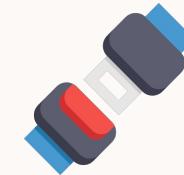
Factors identified in affecting the perception of mobility market



Convenience



Affordability



Safety



Customer Service



Promotion



Accessibility

CORRELATION ANALYSIS

Key Insights



1

Convenience and safety is an important factor for public transport.



2

Despite cost being prevalent, it is not highly correlation to customer satisfaction.



3

Customers seek options that would give them an ease of mind.

CLUSTERING ANALYSIS

Key Insights



Cluster A: **Comfort-Focused**

See importance in comfort
when travelling

Cluster B: **Price-Sensitive**

Feel more strongly about the
price

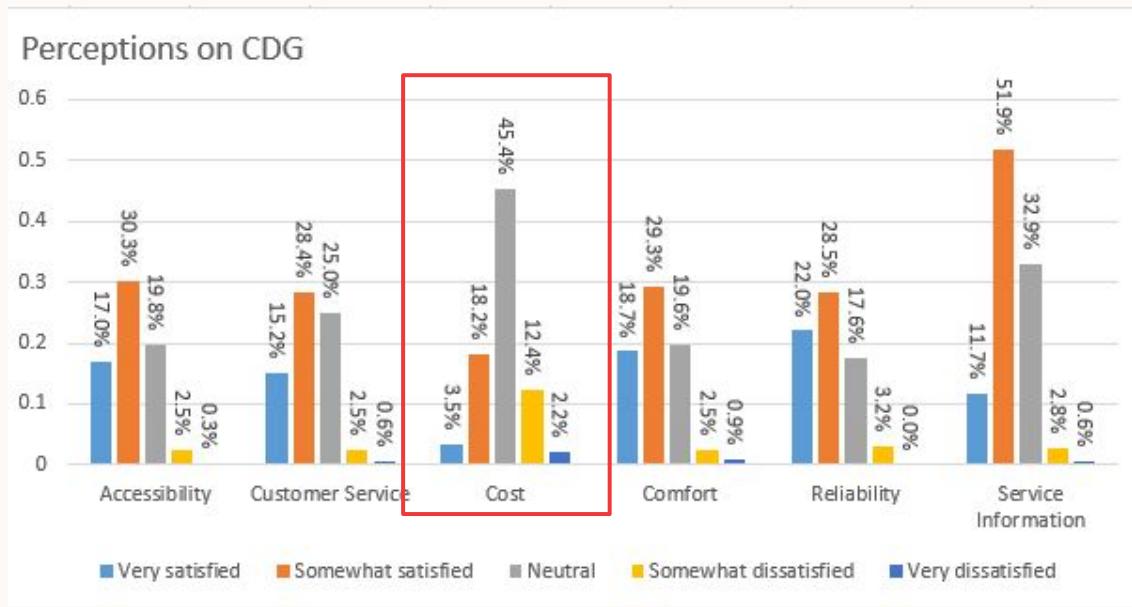
Cluster C: **Experience-Focused**

Promotion, customer
service, accessibility are
important factors to them

CLUSTERING ANALYSIS

Key Insights

Comparing to respondents' perception of CDG's services



GEOSPATIAL ANALYSIS

Key Insights



1

Individual planning area scores are affected by the availability of a transportation service.



2

Effect of new BTOs and MRT stations in a particular planning area can affect its scores.



3

Top Factors that influence level of Satisfaction in each Planning Area:

1. Safety
2. Affordability
3. Service Information



Solution Overview

Approach towards client's solution

APPLICATION GOALS

1



Identify ratings of
different planning area

2



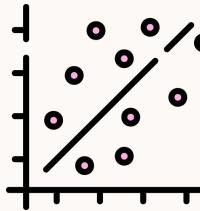
Ranking and Sentiments
of Keywords

3



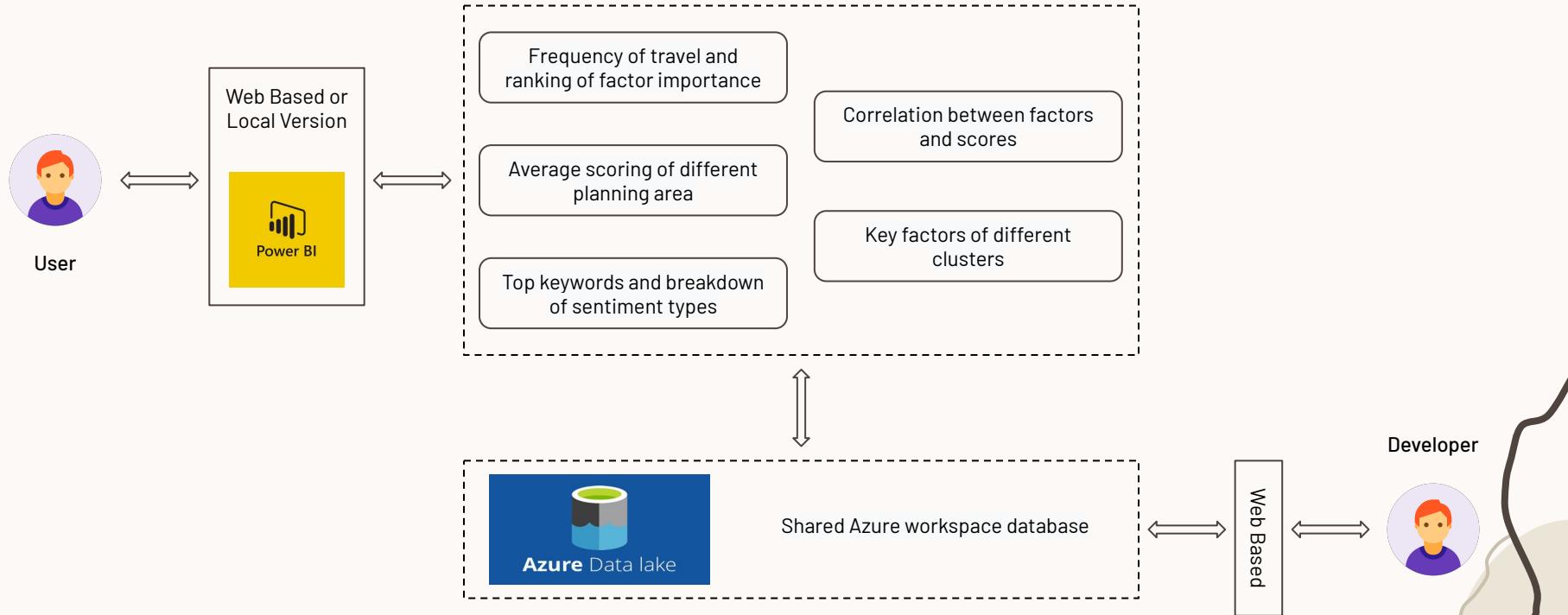
Key factors affecting
different clusters

4

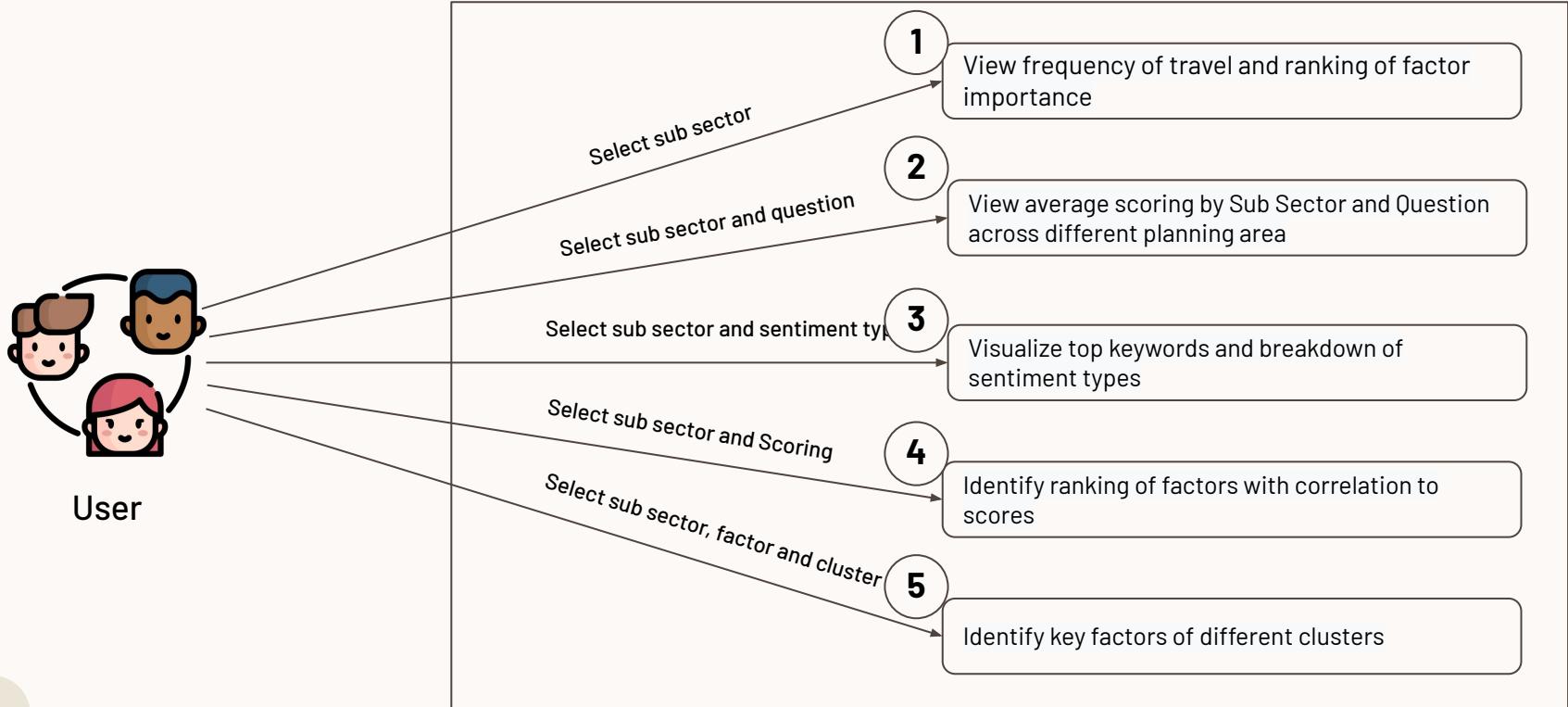


Correlation between
Factors and Scoring

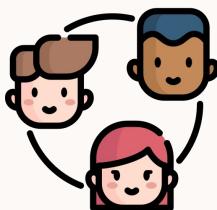
SOLUTION DESIGN



USE CASE DIAGRAM

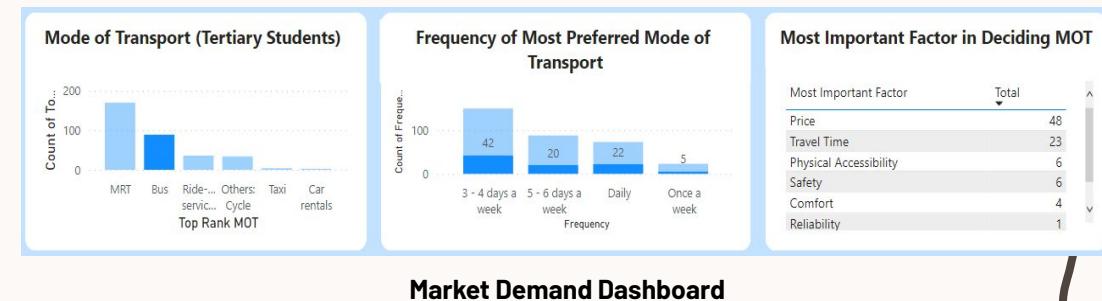


USER SCENARIO 1 - View frequency of travel and ranking of factor importance

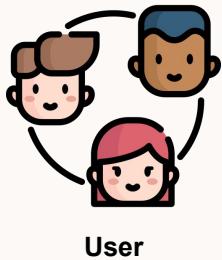


User

1. Click on sub sector from Mode of Transport (Tertiary Students) Barchart
2. View frequency of travel count and ranking of factor importance



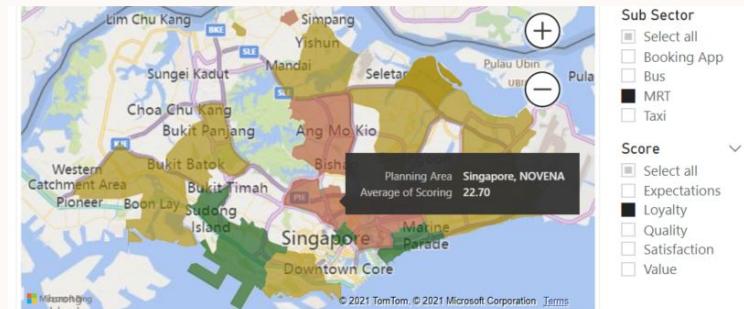
USER SCENARIO 2 - View average scoring by Sub-Sector and Question across different planning area



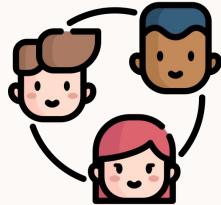
1. Select Sub sector and Question in list to filter



2. View heatmap of question scoring for different planning area. Hover to view exact score.



USER SCENARIO 3 - Visualize top keywords and breakdown of sentiment types

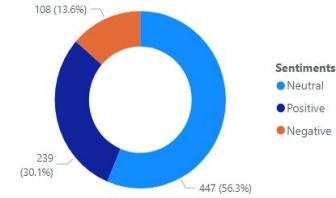


User

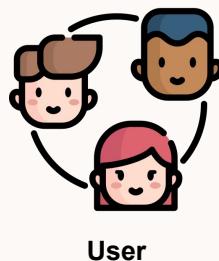
1. Select Sub sector and sentiment types
2. View word cloud to identify top keywords based on size and breakdown of donut chart



Market Demand Dashboard



USER SCENARIO 4 - Identify ranking of factors with correlation to scores



1. Select sub sector and scoring metric to filter
2. View ranking of factors most correlated with chosen filter

CSISG Top Correlated Factors

Factors	Correlation Numbers
Comfort	0.55
Convenience	0.43
Customer Service	0.40
Service Information	0.29
Safety	0.24
Affordability	0.22

Scores

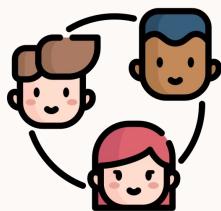
- Customer Expectations Score
- Customer Loyalty User Trust Score
- Customer Satisfaction Score
- Perceived Overall Quality Score
- Perceived Value Score

Sub-Sector

- Booking Application
- Bus
- Taxi
- Train

Market Demand Dashboard

USER SCENARIO 5 - Identify key factors of different clusters



User

1. Select factor and subsector to filter

2. View breakdown of perception on filters for different clusters

3.

Select cluster based on breakdown

4.

View heat map and identify which planning area is most strongly for the factors



Market Demand Dashboard



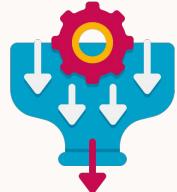
Market Demand Dashboard

Dashboard Demonstration

UAT GOALS

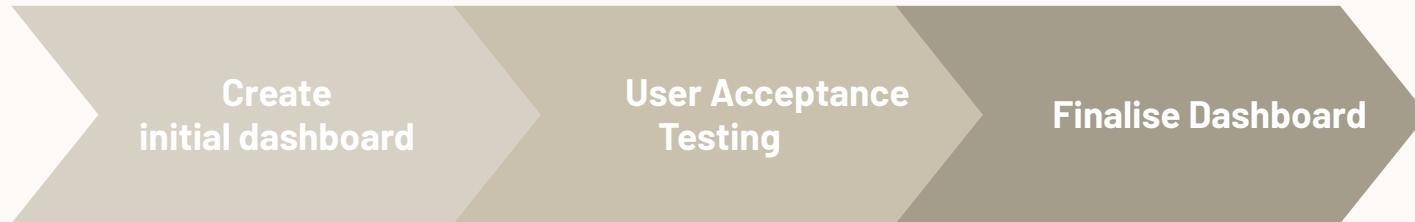


To ensure that dashboard is shows insights that value add to the business user.



To identify and resolve any bottleneck in completing tasks.

OVERVIEW OF UAT



Who 4 members from CDG Data Team

How Due to time constraints and Covid-19 measures, we conducted our UAT online, own-time own-target

What Using metrics such as observations and time taken

Outcome Success

UAT RESULTS

Feedbacks



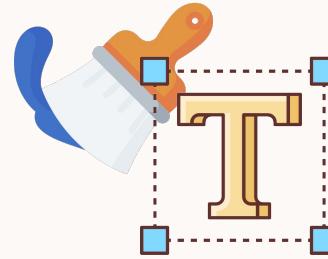
Users are able to conduct the task with ease below the estimated time



Some time required to read the README tab on dashboard



Changes



- Aesthetic changes
- Grammatical errors and typos fixed
- Explanation for the CSISG data
- Simplified the readme tab

POST UAT CHANGES

Before



Gradient Scale Colour

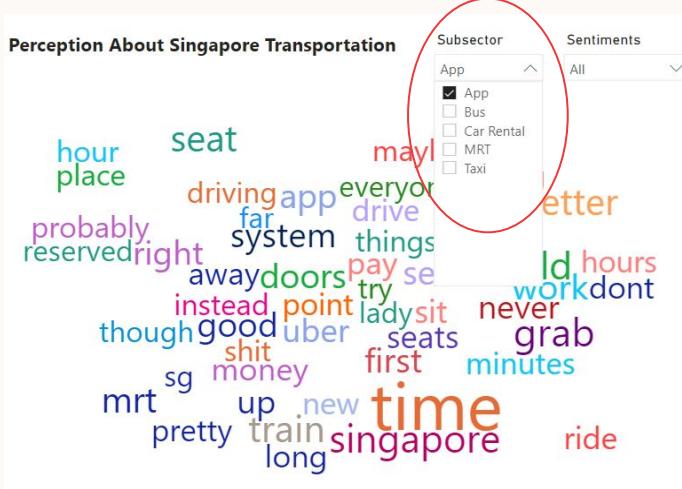
After



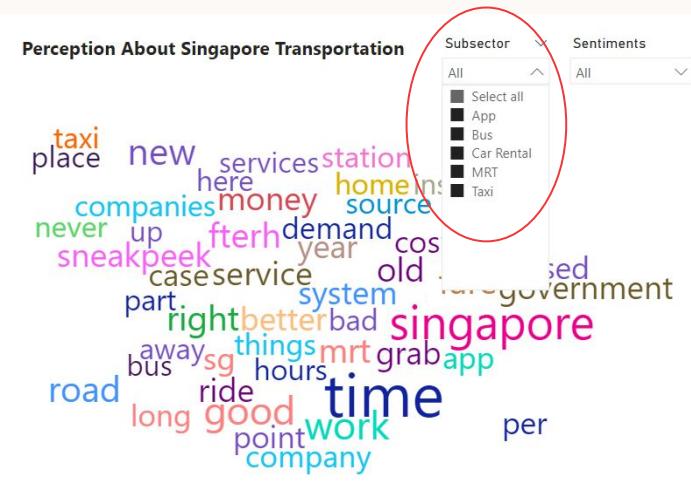
Rule-Based Colour

POST UAT CHANGES

Before



After



Identify ratings of different planning area

POST UAT CHANGES

Before



After



Identify ratings of different planning area



Technical Overview

Tools, technology challenges and data challenges

DATA SOURCES



Survey



Reddit



Customer Satisfaction
Index of Singapore
(CSISG)



Focus Group



data.gov.sg

CHALLENGES



Data Challenges



Insufficient and Irrelevant Data



Readdressing the Problem Statement

Technical Challenges



Difficulties in scraping some data sources



Scraped data eliminated due to irrelevance

Mitigation

Re-evaluate data sources

Timely communication with stakeholders to expedite changes

Relook at data sources and study on the techniques required

Tapped into other resources - survey, CSISG, focus group

TOOLS AND FRAMEWORKS

Programming Languages



Software and Tools



Database



Frameworks and Libraries



API and other Tools

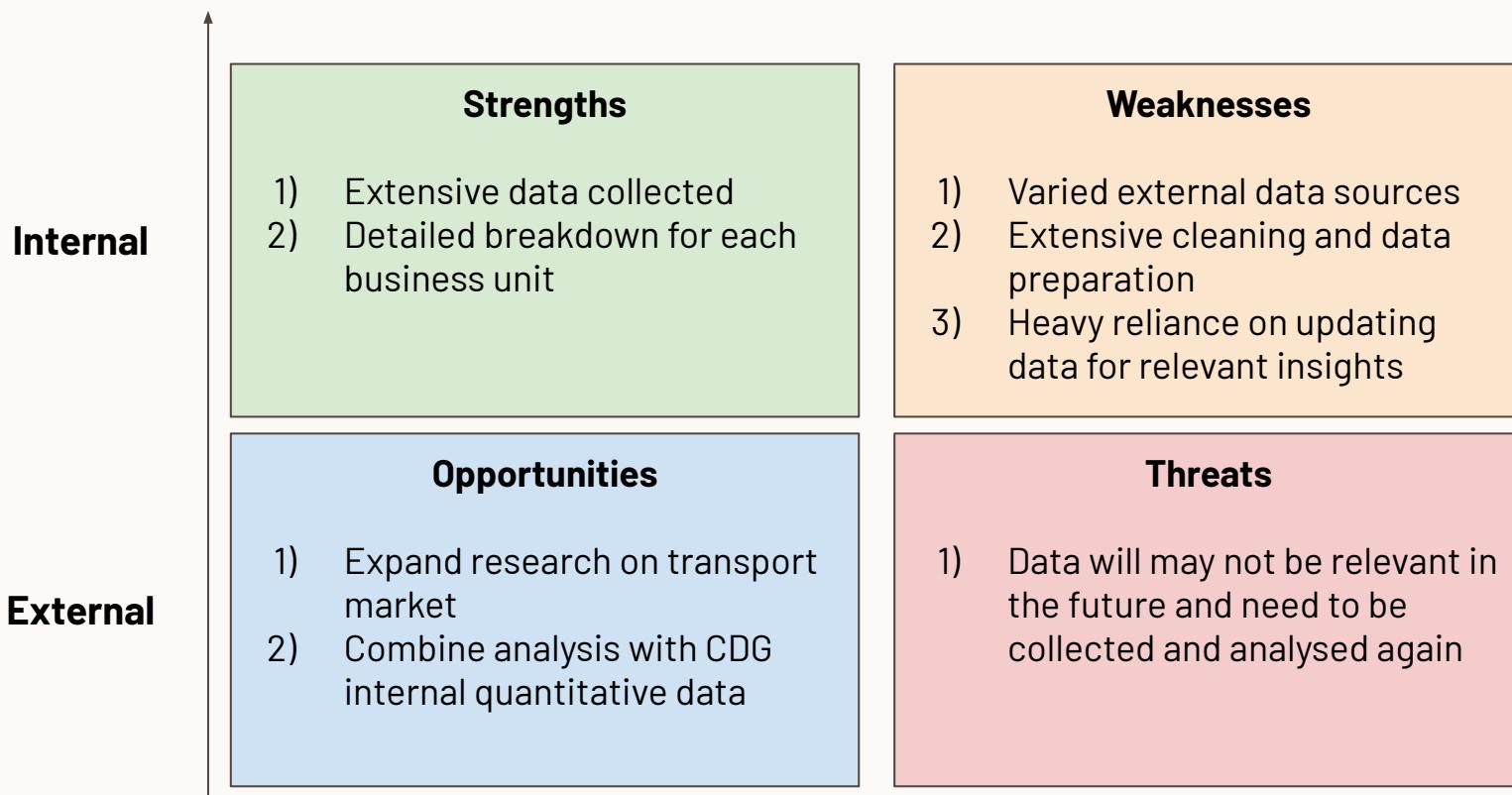




Gap Analysis

Analysis of the project's gaps

SWOT ANALYSIS





Moving Forward

FUTURE WORK

Automation of manual processes

- Value add to the process of data collection in scraping data
- Data will refresh frequently and easily allowing more time to examine insights more accurately

Direct Data Pipeline

- Microsoft Power Automate to create automated workflows, directly funnel survey data into dashboard
- Real-time data and insights

Utilisation of more complex machine learning algorithms

- Run modern neural network classifications
- Hyperparameter tuning to determine the right combination of complex algorithms
- Maximise models' performance and more accurate insights

HANOVER

PowerBI Dashboard

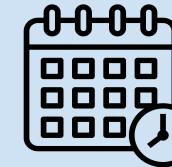
Source Code

Report of research, insights and recommendations

User Guide for analyses and instructions of deliverables

Poster

Video



Expect to handover within 1 week from today



Conclusion

Recommendation and Project Takeaway

OVERALL KEY INSIGHTS

1

Safety is a prevalent factor

3

Low pick-up rate of car rental service

2

Relatively sensitive to the cost of service

4

Profiling of tertiary students

RECOMMENDATIONS



Social listening to understand the needs and priorities of our target audience



Review prices across each business units through pricing bundles, as well as reevaluate the frequency of transport services



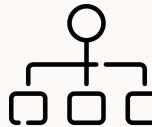
Optimise marketing strategies to cater to the characteristics of the different cluster



Safety should be an important factor of consideration for all subsectors

PROJECT TAKEAWAYS

Client



A systematic framework to follow and make changes for future projects



All in one analytical dashboard to generate meaningful insights and conclusions

Team



Being adaptable to scope changes in a real-world context



Increase proficiency in both technical and soft skills

Thank you!

Q&A

