Satisfaction of active mobility program students with the Binus Anggrek cafeteria

I Kadek Defa Danuarta

Backgrounds

We are a team of active students in the mobility program. My team and I have had several new experiences, starting with small things in campus life, one of which is shopping at the canteen. My colleagues come from different campuses, starting from Binus Semarang, Binus Bandung, and myself from Binus Malang.

In this project, I will attempt to analyze **how active students in the student mobility program perceive the service, cleanliness, and food quality at the Binus Anggrek cafeteria**, and **analyze the factors influencing satisfaction based on questionnaire data.**

Datasets

This data was obtained from an internal survey on student satisfaction levels regarding service, cleanliness, and food products at the Binus Anggrek Campus cafeteria. The survey was conducted anonymously using a Likert scale in bahasa language.

DISCLAIMER
The data has been anonymized by removing personal identity information such as names, email addresses, and student ID numbers in accordance with personal data protection principles. *The data will be extracted to a .csv file format in every type of changes in the process*

Data Cleaning & Preprocessing

Identifies Missing Value and The Causes

```
## Missing values is founded in:
```

```
## major : 6 missing value(s)
## home_campus : 4 missing value(s)
```

The main reason for the missing values, especially for *Home Campus*, is that the respondents are not active participants in the mobility program. Since the purpose is to measure the satisfaction of **students who are actively involved in the mobility program**, I decided to remove these entries with missing *Home Campus* values.

For column **Jurusan**, I imputed the missing values using the most frequently occurring major. This approach is intended to preserve useful responses, as the remaining data may still contribute valuable insights for the analysis.

```
## Missing values after deletion and imputation:
```

```
## major: 0
```

```
## home_campus: 0
```

From the results above it shows a successful attemp to handle missing values

Handle Inconsistent Values in Likert Scale

To anticipate inconsistent values, I'll change all the likert categories to a lowercase (in case there are some miss type in the questionnaire form)

```
## Before mutation: Setuju, Sangat Setuju, Tidak merasakan, Sangat Tidak Setuju
```

```
## After mutation: setuju, sangat setuju, tidak merasakan, sangat tidak setuju
```

This process is to anticipate errors in encoding process, since the inconsistent value is completely handled, encoding is needed to transform the likert category to a numeric form of scale.

Separating Date, Timestamp, and Durations

```
## # A tibble: 6 x 2
##   start_time          completion_time
##   <dttm>              <dttm>
## 1 2025-04-05 18:35:56 2025-04-05 18:37:57
## 2 2025-04-05 18:38:42 2025-04-05 18:42:44
## 3 2025-04-05 18:36:07 2025-04-05 18:43:53
## 4 2025-04-05 18:39:48 2025-04-05 18:44:13
## 5 2025-04-05 19:11:49 2025-04-05 19:16:21
## 6 2025-04-05 21:32:59 2025-04-05 21:35:02
```

By default, the tool used to gather questionnaire results has a time stamp about starting time to a completion time. Where in this case I will parse them to an identical column. In this case I'll separate the date and time into an individual column.

```
## # A tibble: 6 x 6
##      ID major            date       start_time completion_time duration
##   <dbl> <chr>            <date>     <chr>      <chr>              <dbl>
## 1    12 digital business 2025-04-05 18:35:56   18:37:57            2.02
## 2    14 computer science 2025-04-05 18:38:42   18:42:44            4.03
## 3    15 computer science 2025-04-05 18:36:07   18:43:53            7.77
## 4    16 computer science 2025-04-05 18:39:48   18:44:13            4.42
## 5    17 computer science 2025-04-05 19:11:49   19:16:21            4.53
## 6    18 computer science 2025-04-05 21:32:59   21:35:02            2.05
```

Encoding Likert Scale and frequency

```
## Unique value for every likert scale:  setuju, sangat setuju, sangat tidak setuju, tidak setuju, tidal
```

The result above are unique values. Where in this case, every likert scales will be encoded manually to a range number from one(1) to five(5).

```
## Unique value for every likert scale (after encoding):  4, 5, 1, 2, 3
```

New Column for Further Analysis

Average/Mean Satisfaction Score
Here I make a new column to support further analysis and prediction tasks. This value is calculated as the **average of every encoded likert scale**. Why is this important? it provides a single quantitative measure that summarizes overall student satisfaction. It allows for easier comparison between respondents and can be used as a target variable in predictive modeling or segmentation analysis.

```
## # A tibble: 6 x 6
##      ID major      mean_services_score mean_cleanliness_score mean_product_score
##   <dbl> <chr>                    <dbl>                  <dbl>              <dbl>
## 1    12 digital b~                   4                   4                  4
## 2    14 computer ~                4.57                3.71               3.67
## 3    15 computer ~                4.14                4.29               4.33
## 4    16 computer ~                4.43                   3               4.83
## 5    17 computer ~                   4                2.57               4.17
## 6    18 computer ~                   4                   3                  3
## # i 1 more variable: mean_satisfaction_score <dbl>
```

Average/Mean Score in Every aspects
To perform an in-depth analysis of the performance of each metric measured such as services, cleanliness, and product, I created a new column to store the average of each question to measure respondents' satisfaction with the aspects

Analysis

Distribution of Duration

The first analysis is to see the distribution of duration. This graph might help to evaluate respondents who fill out the questionnaire too quickly or too slowly, which indicates their seriousness in answering the questionnaire.

Distribution of Duration (minutes)
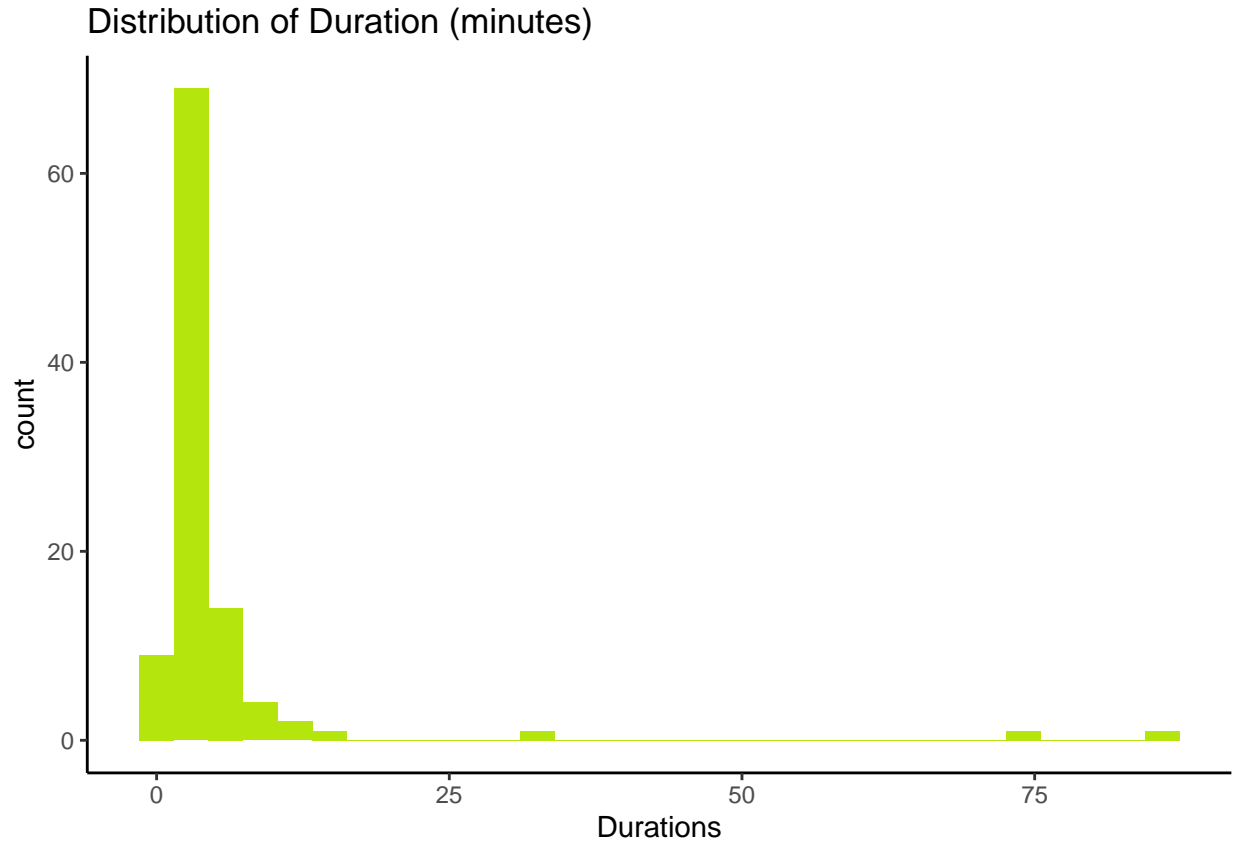
Based on the histogram above, respondents are more likely to complete the questionnaire in about 2 minutes. This data shows that respondents answered the questionnaire fairly quickly.

At 0 minutes, this could be caused by some respondents answered the questionnaire very quickly and might not read the questions (not taking it seriously). Further analysis will be explained through the descriptive statistics below.

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.520   2.277   3.065   5.427   4.082  86.400
```

The shortest duration for respondent to fill questionnaire is 0.520 minutes.

Canteen Visitors Frequency

In this analysis, I'll try to measure visitation rate with satisfaction. In order to categorize the satisfaction level, I'll use discretization or binning to make a category out of mean satisfaction score. The mapping of this process are as follows:

(MSS = Mean Satisfaction Score)

- MSS less or equal to 1 = "Bad"

- MSS greater than 1 AND MSS less or equal to 3 = "Quite Bad"

- MSS greater than 3 AND MSS less than 5 = "Good"

- MSS equal to 5 = "Excellent"

Graph View   After discretization or binning process, the analysis will be utilizing a plot to show the
Amount of Canteen Visitors by Frequent Visiting and Satisfaction Category

Canteen Visitors by Frequent Visiting and Satisfaction Category
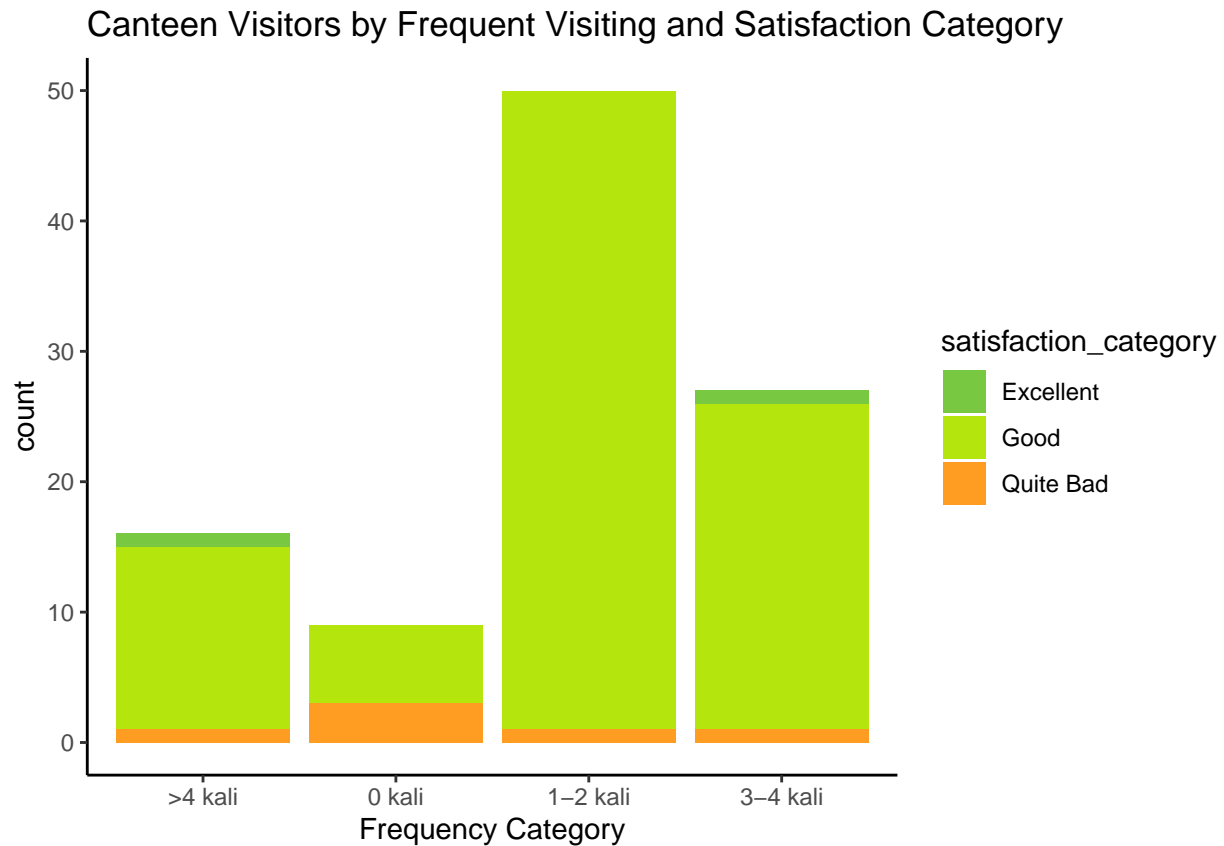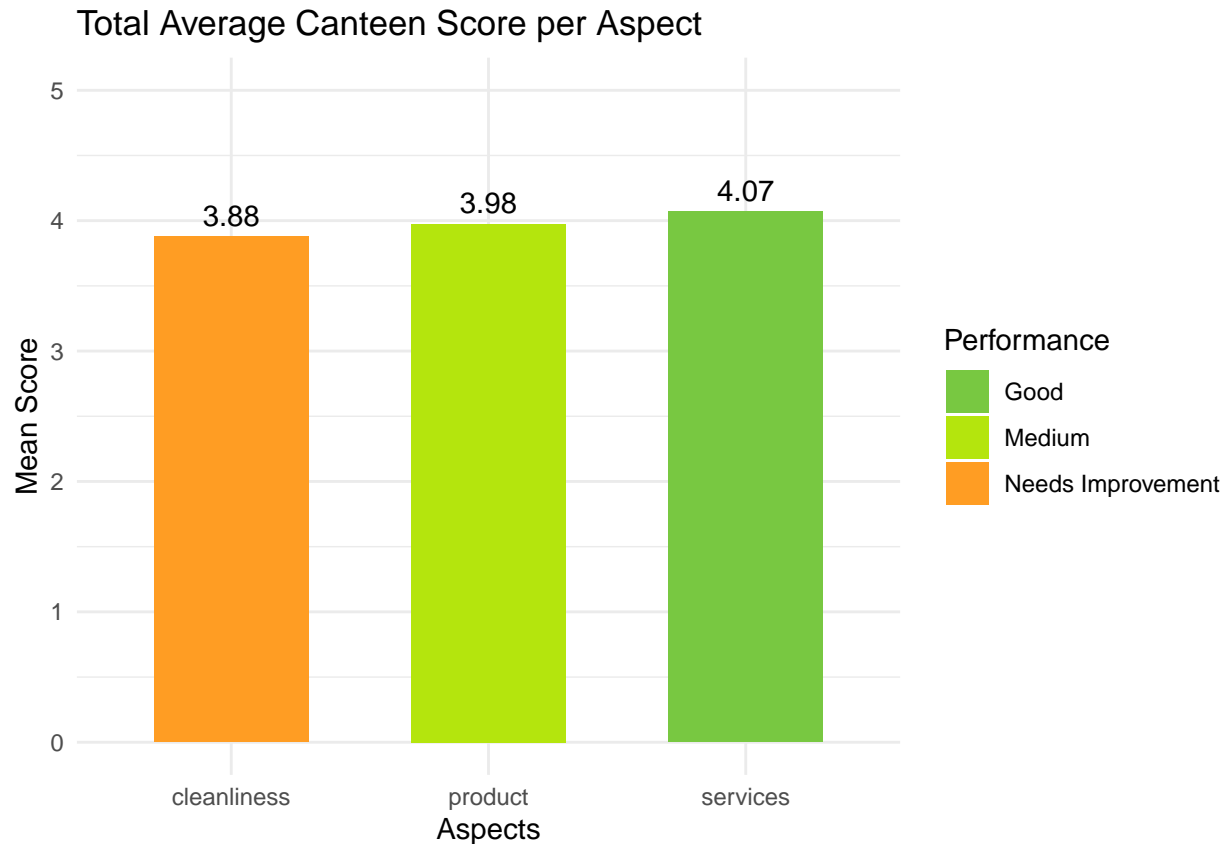
Explanation:

- The visitors mostly come once until two times a week, meaning that the canteen is not a frequent place to visit. The overall satisfaction level is on "good" condition with 2 respond show an excellent satisfaction level.

- Respondent who chose "0 kali" or rarely go to canteen showing the most "quite bad" satisfaction level among others. This might indicated by some people that had a unfulfilled expectations.

Table View

```
## # A tibble: 6 x 4
##      ID major            mean_satisfaction_score satisfaction_category
##   <dbl> <chr>                              <dbl> <chr>
## 1    12 digital business                       4 Good
## 2    14 computer science                       4 Good
## 3    15 computer science                    4.25 Good
## 4    16 computer science                    4.05 Good
## 5    17 computer science                    3.55 Good
## 6    18 computer science                    3.35 Good
```

Performances in every aspects

This analysis is conducted to know the performance of every aspects that measured. The aspects that are measured such as, service, cleanliness, and product. The graph below shows the total average of the average performance scores for each aspect. This graph is expected to provide an overview of aspects that can be improved in the Binus Anggrek cafeteria.


Total Average Canteen Score per Aspect

Explanation

- Based on the graph above, respondents gave **fairly good responses to the service**, as indicated by the average total score falling into the "good" category.

- This was followed by the performance of the canteen food, which received a "medium" rating, with **room for improvement in terms of product quality and price.**

- Finally, the aspect of **cleanliness requires significant improvement** in terms of cleanliness and access to washing facilities.

Seeing the Weaknesses of each Aspect

In this analysis, I will focus on aspects that have a "Need Improvements" status which referring to Product and Cleanliness aspect. Here, I will create a descriptive statistical table to dig deeper into the areas that are weaknesses that greatly affect the average student satisfaction score for the Binus Anggrek cafeteria.

In this process, it'll require binning technique to categorize each variables based on their average scores. The Binning process will be mapped as follows:

- mean less than 3.5 = "Needs Improvement"

- mean less than 4 = "Medium"

- mean greater than or equal 4 = "Good"

Average Score per Variable on Cleanliness Aspect
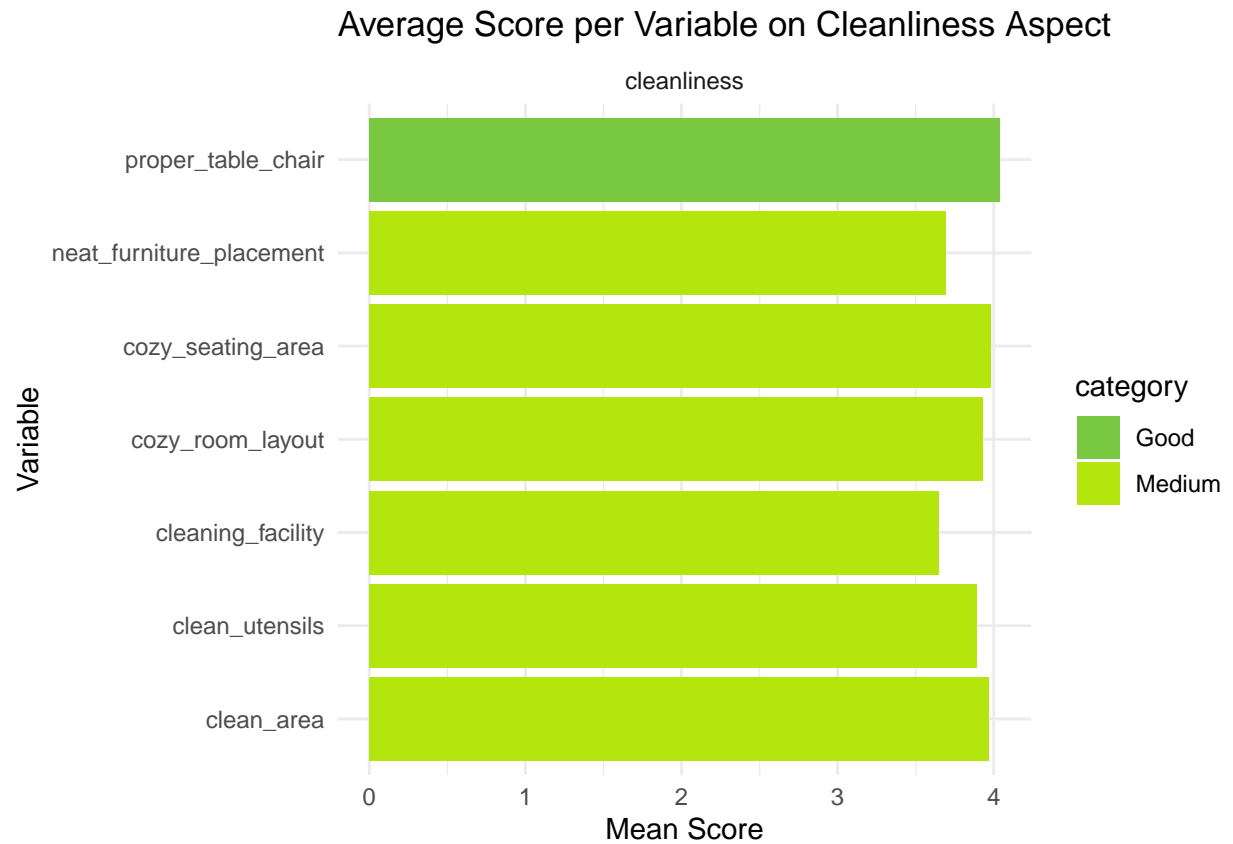
Graph View

Cleanliness   The average of `cleaning_facility` variable have the lowest results such as 3,62 (if rounded by 2), followed by `neat_furniture_placement` with result of the average is 3,69. These both variable are the weak point for cleanliness aspect. This weaknesses also affecting the total average score of cleanliness (could be found at "Analysis" tab) to be the lowest among all existing aspects.

Both of those variable have a high standard deviation (sd > 0,5) which indicate that the distribution was quite wide among the respondents' answers. Moreover, there was a change in the location of the cafeteria during the collection of the questionnaire results, it might affect the wide response that has mentioned before.

Average Score per Variable on Product Aspect
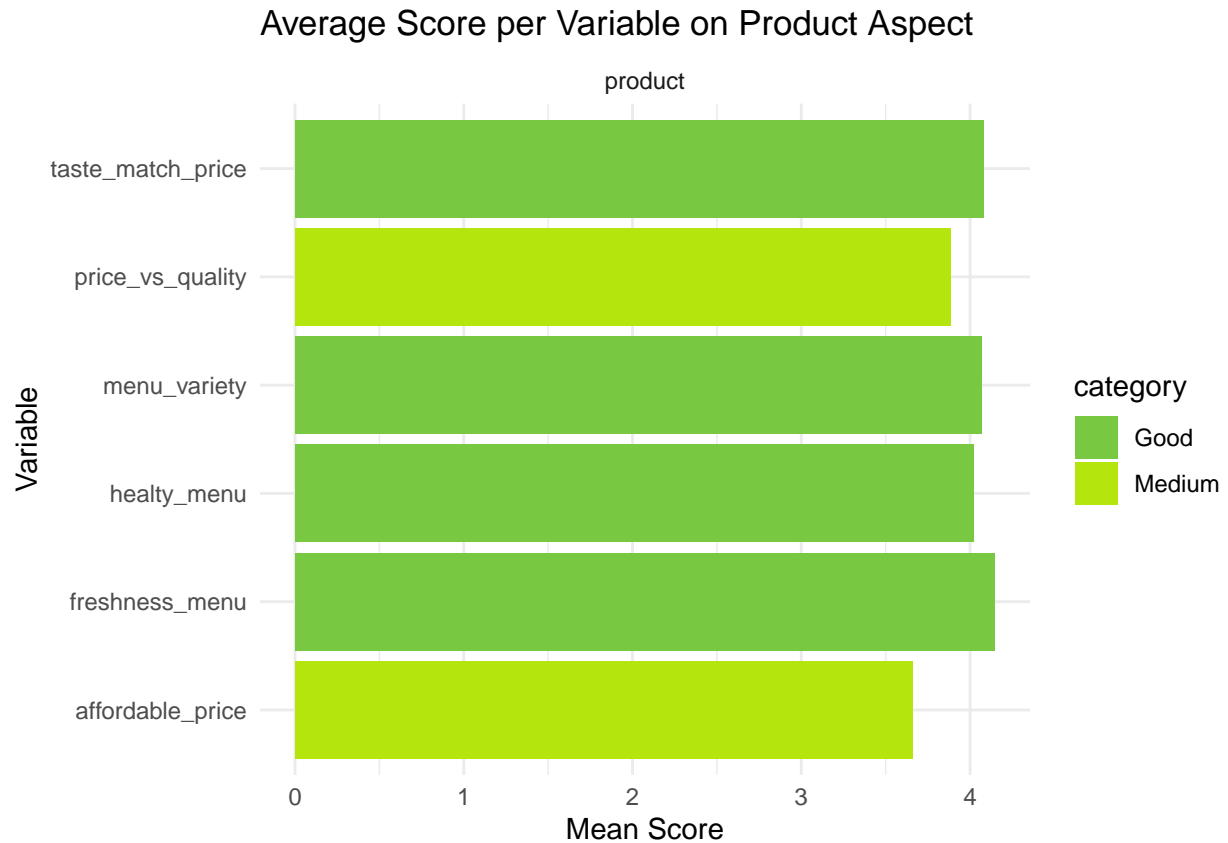
Product   In Product aspect, there are two different variables that fell into "Medium" category, such as `affordable_price` with the average of 3,65 and `price_vs_quality` with an average of 3,88. The `affordable_price` variable had a high standard deviation (~0,93), this indicates a fairly wide spread among respondents' answers and may be influenced by various segments available.

# Average Score per Variable on Service Aspect
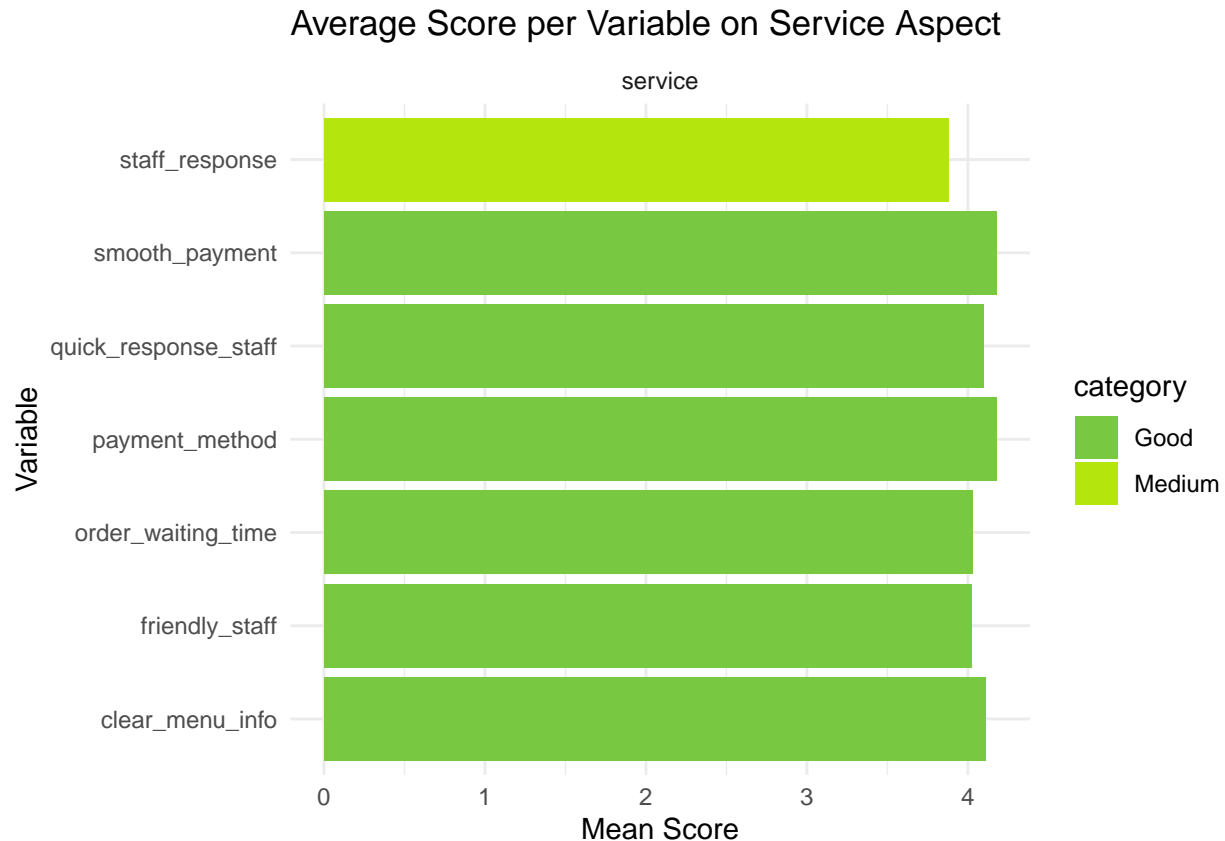
service

Services Finally, most respondents stated that service aspects fell into the "Good" category, as seen from the average of each variable, which was greater than 4 ("Agree" on the Likert scale) and the consistency of respondents (median) answering 4 on most service aspects. However, despite the ratings falling into the "Good" category, `staff_response` falls into the "Medium" category, as not all respondents felt that the staff in the cafeteria were responsive to questions or complaints.

Table View

```
## # A tibble: 6 x 8
## # Groups:   aspect [1]
##   aspect      variable                 mean median    sd   max   min category
##   <chr>       <chr>                   <dbl>  <dbl> <dbl> <int> <int> <chr>
## 1 cleanliness clean_area               3.97      4 0.696     5     1 Medium
## 2 cleanliness clean_utensils           3.89      4 0.643     5     2 Medium
## 3 cleanliness cleaning_facility        3.65      4 0.961     5     1 Medium
## 4 cleanliness cozy_room_layout         3.93      4 0.707     5     1 Medium
## 5 cleanliness cozy_seating_area        3.98      4 0.832     5     1 Medium
## 6 cleanliness neat_furniture_placement 3.70      4 0.876     5     2 Medium
```

Segmentation Based on Home Campus

This analysis was conducted to gain insight into the responses of respondents from various campuses regarding three aspects that influence student satisfaction with the Binus Anggrek cafeteria.

Satisfaction per Home Campus Segmentation
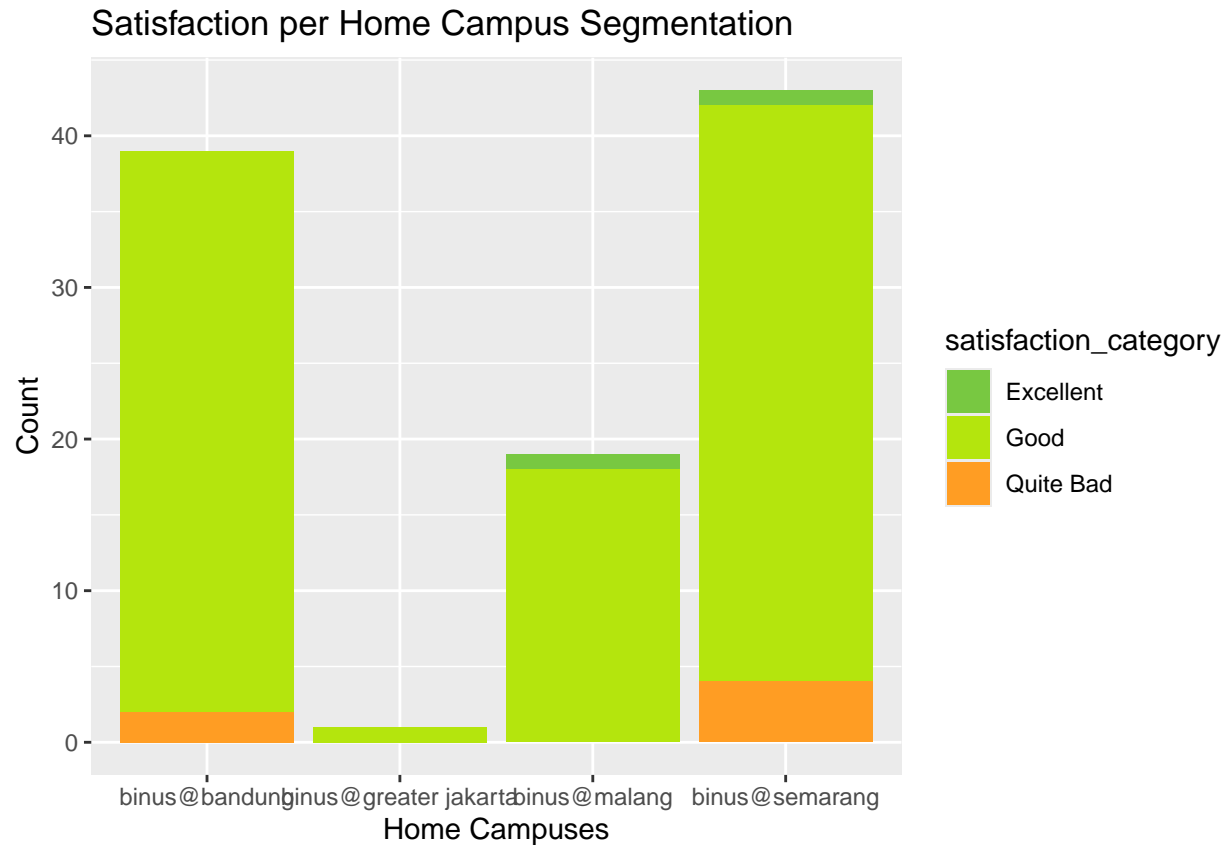
Explanations:

Based on home campus segmentation, the majority of students from **Binus Bandung and Binus Semarang have a satisfaction category classified as "Good,"** with a few of them falling into the "Excellent" category. On the other hand, some respondents from **Binus Semarang fall into the "Quite Bad"** review category, which may be due to their lowered expectations regarding the Binus Anggrek canteen. Based on the above results, although satisfaction levels are relatively stable, there is **room for improvement to make the experience truly exceptional.**

Average Score per Aspect Grouped by Home Campus

This analysis was conducted to determine the level of satisfaction of respondents from different home campuses with regard to the three aspects of the cafeteria.

Average Score per Aspect by Home Campus
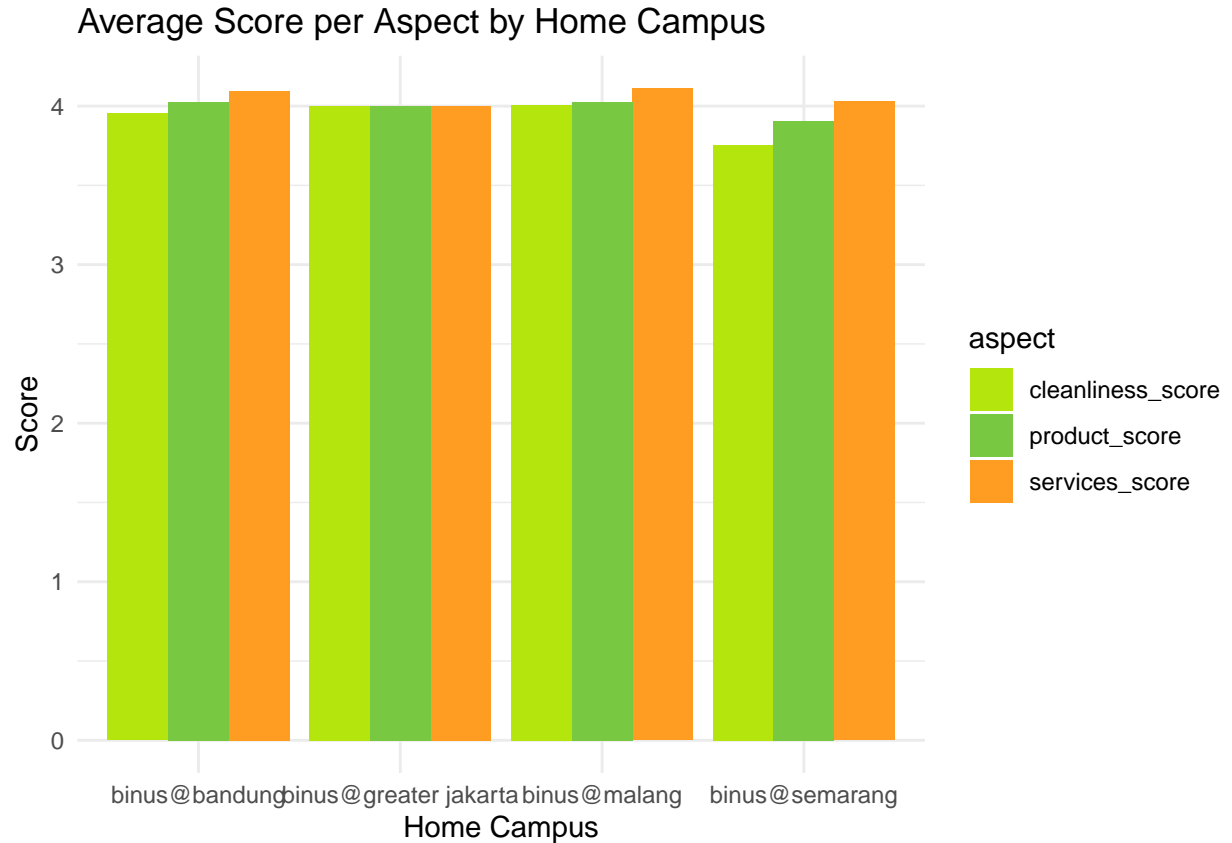
Graph View

Main Insight:

1. **The Service aspect** (services_score) consistently received the highest score across all campuses, with the highest score belonging to Binus Malang (4.112) and the lowest to Binus Greater Jakarta (4.000).

2. **Cleanliness** (cleanliness_score) shows the greatest difference between campuses:
   2.1. *Highest score:* Binus Malang (4.008)
   2.2. *Lowest score:* Binus Semarang (3.754), this score also caused this campus to be the only one in the "Needs Improvement" category in the previous analysis.

3. **Products** (product_score) tend to be in the middle, relatively consistent with small differences between campuses (~0.1).

Explanation for Binus Greater Jakarta: The graph above shows a comparison of scores for each aspect of the Binus Anggrek cafeteria assessment, grouped by home campuses. Some of the data above cannot be confirmed for accuracy, particularly for respondents from the Binus Greater Jakarta campus, as there was only one respondent who scored a 4 on all three aspects. Therefore, the results of respondent from greater jakarta should not be used as a basis for decision-making.

Table View

```
## # A tibble: 4 x 6
##   home_campus          responses services_score cleanliness_score product_score
##   <chr>                    <int>          <dbl>             <dbl>         <dbl>
```

```
## 1 binus@bandung                   39         4.10          3.95          4.03
## 2 binus@greater jakarta            1         4             4             4
## 3 binus@malang                    19         4.11          4.01          4.03
## 4 binus@semarang                  43         4.03          3.75          3.91
## # i 1 more variable: overall_score <dbl>
```
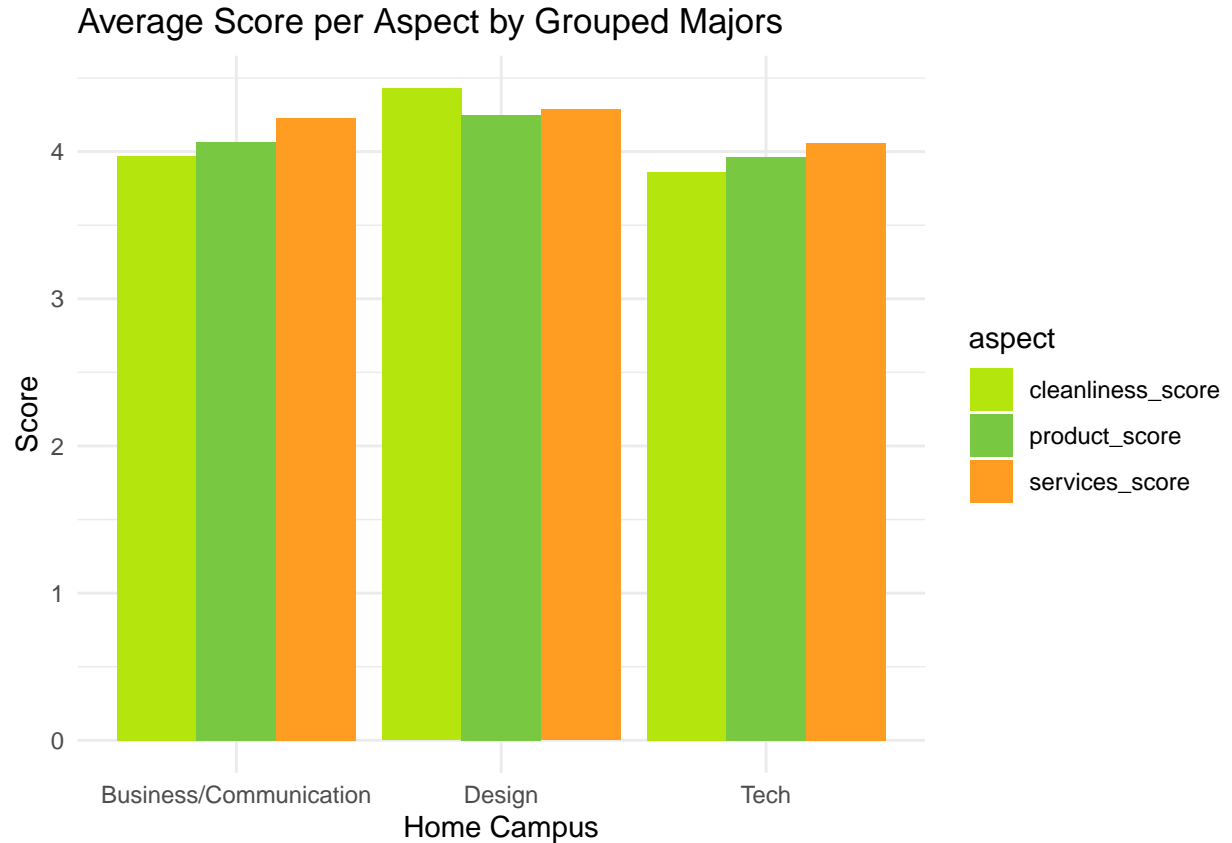
Segmentation Based on Major

This analysis was conducted to gain insight into the average satisfaction scores of students grouped by major. Therefore I'll check the unique values in case there were so many majors are in the datasets.

```
##  [1] "digital business"
##  [2] "computer science"
##  [3] "public relations"
##  [4] "communication - marketing communication"
##  [5] "information systems"
##  [6] "data science"
##  [7] "computer science - software engineering"
##  [8] "artificial intelligence"
##  [9] "cyber security"
## [10] "game application and technology"
## [11] "interior design"
## [12] "computer engineering"
```

There are 12 unique values in major that might require to turn them into a broad categories. For example:

- **Tech / IT:** Computer Science, Software Engineering, AI, Cyber Security, Computer Engineering, Data Science, Information Systems

- **Design / Creative:** Interior Design, Game Application

- **Business / Comm:** Digital Business, Marketing Comm, Public Relations

Average Score per Aspect by Grouped Majors

Graph View

Insights & Explanations:
Based on the segmentation of grouped majors, students from the Tech field dominate the respondent population (n = 95), making the scores from this group the most representative reference.

On the other hand, the Design and Business/Communication groups showed relatively higher satisfaction scores. However, due to the very small number of respondents (n = 2 and n = 5), these results are not sufficiently robust to be generalized. This finding underscores **the importance of balanced representation in surveys to ensure that the analysis reflects the entire student population.**

Table View

```
## # A tibble: 3 x 6
##   major_group          responses services_score cleanliness_score product_score
##   <chr>                    <int>          <dbl>             <dbl>         <dbl>
## 1 Business/Communicati~        5           4.23              3.97          4.07
## 2 Design                       2           4.28              4.43          4.25
## 3 Tech                        95           4.06              3.86          3.96
## # i 1 more variable: overall_score <dbl>
```

Conclusions

Analisa menunjukan bahwa tingkat kepuasan mahasiswa memiliki variasi yang luas di setiap aspeknya. Pada analisa sebelumnya menyatakan bahwa kebanyakan responden menjawab kuisioner dengan kurun waktu 2 menit, waktu ter singkat adalah

Summary of Key Findings