**Comprehensive Analysis and Dietary Strategies with Tableau: A College Food Choices Case Study**

"Comprehensive Analysis and Dietary Strategies with Tableau: A College Food Choices Case Study" is an innovative project aimed at revolutionizing how dietary data among college students is visualized and utilized to drive informed decision-making and enhance student health and academic performance. In today's educational environment, it's crucial to have access to comprehensive insights into dietary trends, nutritional habits, and health impacts to empower stakeholders with actionable information.

This project seeks to create a dynamic and intuitive platform using Tableau, where data from various aspects of student diets, exercise habits, and health perceptions can be transformed into interactive visualizations and insightful analytics. By leveraging Tableau's capabilities effectively, the "Enhancing Dietary Strategies" project aims to empower educational institutions with actionable insights, foster data-driven decision-making, and drive student well-being by facilitating a deeper understanding of dietary dynamics and promoting evidence-based nutritional strategies.

Scenarios:

Scenario 1: Monitoring Nutritional Intake

In a real-time scenario, imagine receiving an alert indicating a concerning trend in nutritional intake among students, such as a significant decrease in fruit and vegetable consumption. Using the College Food Choices data, we can quickly assess the extent and potential impact of this trend, identify contributing factors, and deploy immediate interventions to encourage healthier eating habits. Whether it’s through targeted awareness campaigns, adjustments in cafeteria menus, or personalized dietary advice, real-time analysis enables rapid decision-making and proactive measures to promote student health.

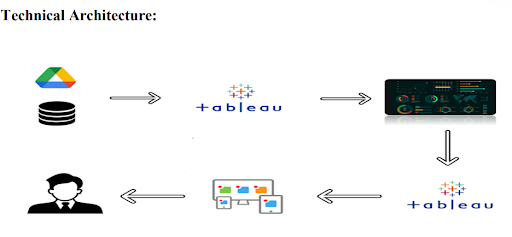
Scenario 2: Addressing Dietary Deficiencies

In the event of identifying widespread dietary deficiencies, such as low vitamin intake or high consumption of unhealthy snacks, real-time access to College Food Choices data enables swift response and management. University health services and nutritionists can utilize the dataset to gather crucial information about the deficiencies, including their prevalence, potential health impacts, and affected student demographics. By leveraging real-time analytics, they can coordinate health promotion efforts, allocate resources effectively, and implement educational programs to address the deficiencies and ensure the nutritional well-being of all students.

Scenario 3: Predictive Analysis and Personalized Nutrition Plans

Leveraging predictive analytics capabilities, College Food Choices empowers universities to anticipate and prevent potential health issues related to poor dietary habits. By analyzing historical data and identifying predictive indicators, health professionals can proactively address nutritional gaps, unhealthy eating patterns, and other risk factors that could lead to health problems. Real-time monitoring of dietary choices, meal consumption patterns, and nutritional intake enables timely interventions, personalized nutrition plans, and continuous support to encourage long-term healthy eating habits among students.

**Technical Architecture**



**Project Flow**

To accomplish this, we have to complete all the activities listed below,

Data Collection & Extraction from Database

o Collect the dataset,

Data Preparation

o Prepare the Data for Visualization

Data Visualizations

o No of Unique Visualizations

Dashboard

o Responsive and Design of Dashboard

Story

 o No of Scenes of Story

Performance Testing

o Amount of Data Loaded

o Utilization of Data Filters

o No of Calculation Fields

o No of Visualizations/ Graphs

Web Integration

o Dashboard and Story embed with UI With Flask

Project Demonstration & Documentation

o Record explanation Video for project end to end solution

o Project Documentation-Step by step project development procedure

**Data Collection & Extraction from Database**

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

**Downloading the dataset**

Please use the link to download the dataset:[link](https://www.kaggle.com/datasets/borapajo/food-choices?select=food_coded.csv): <https://www.kaggle.com/datasets/borapajo/food-choices?select=food_coded.csv>

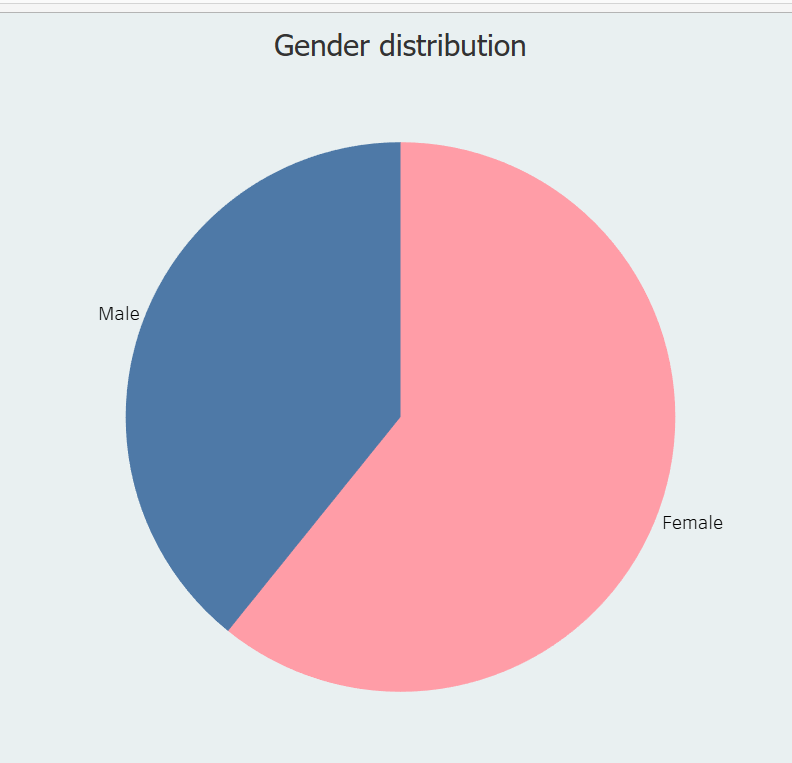
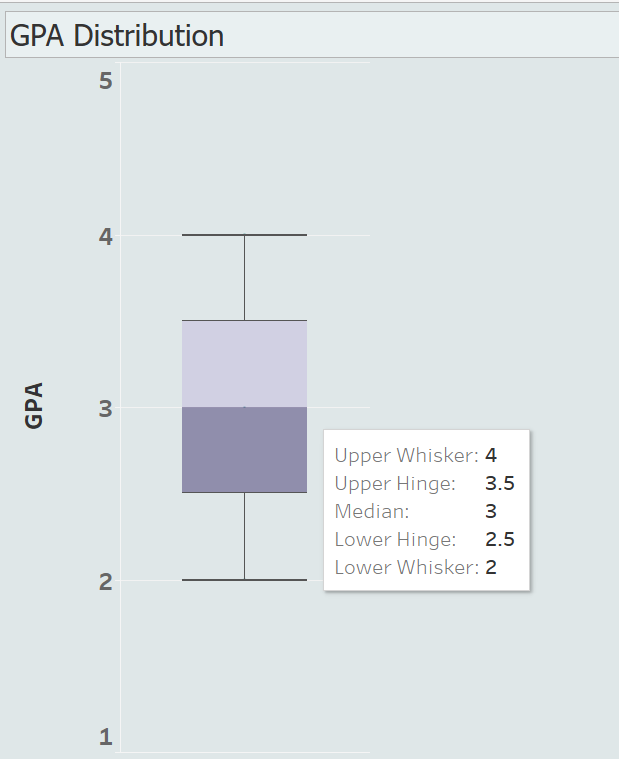
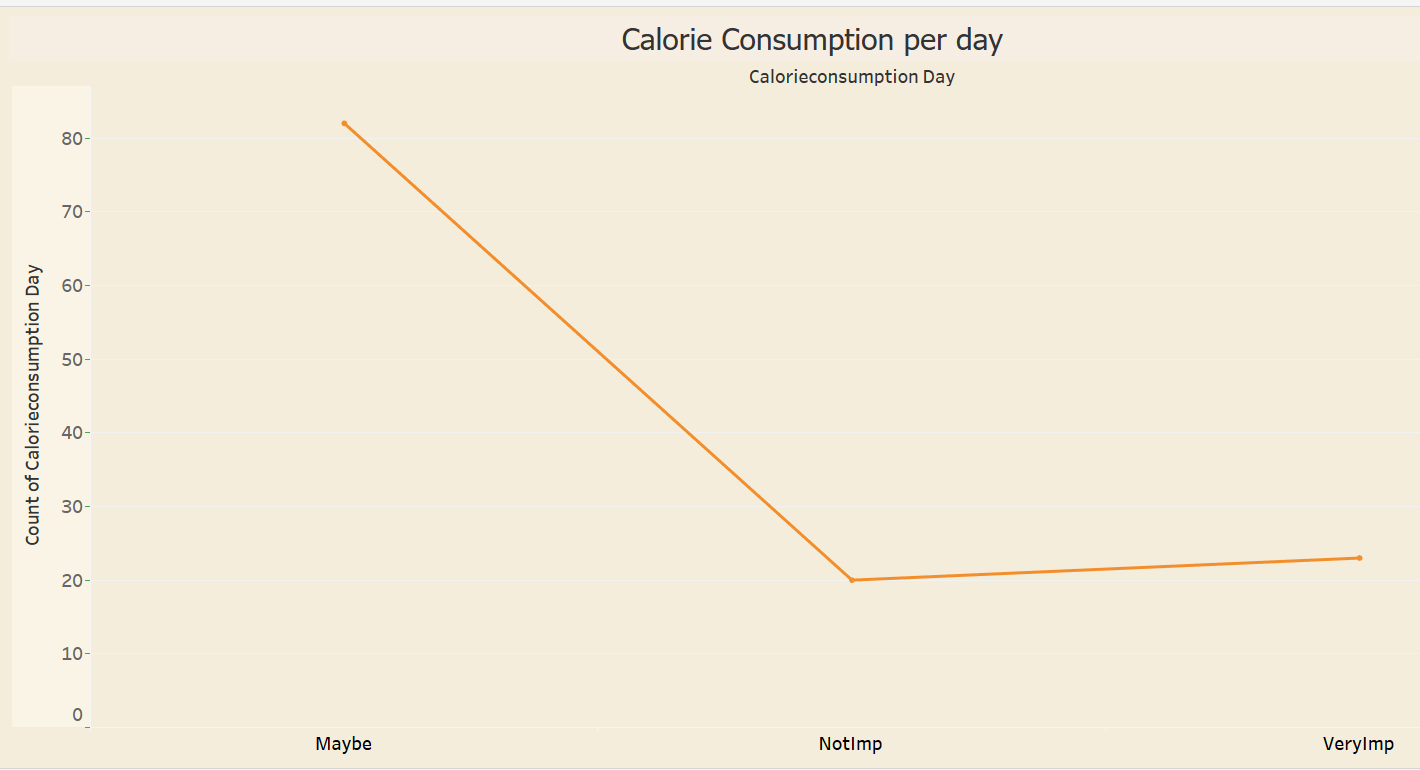
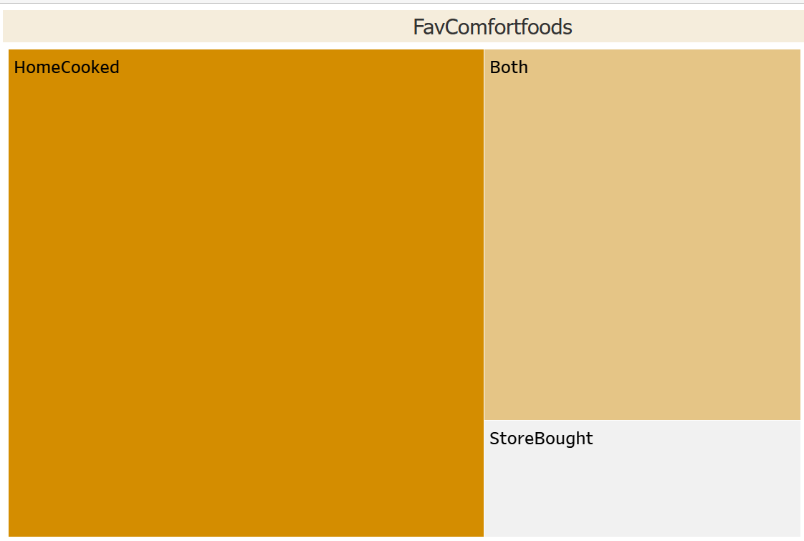
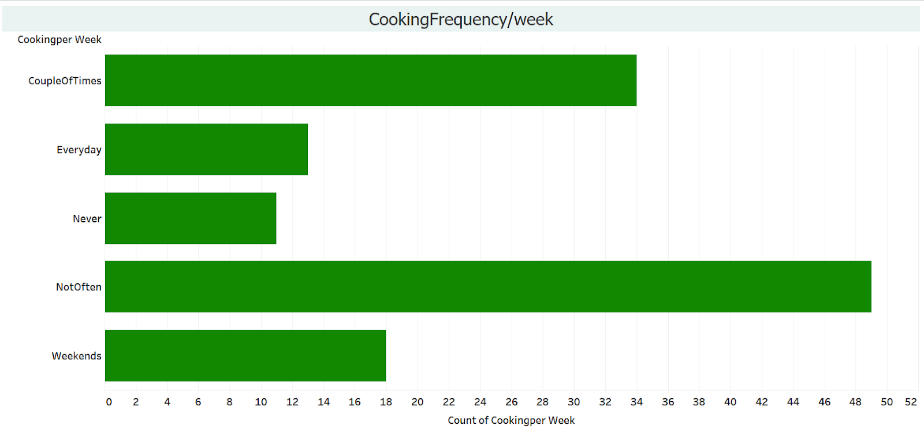
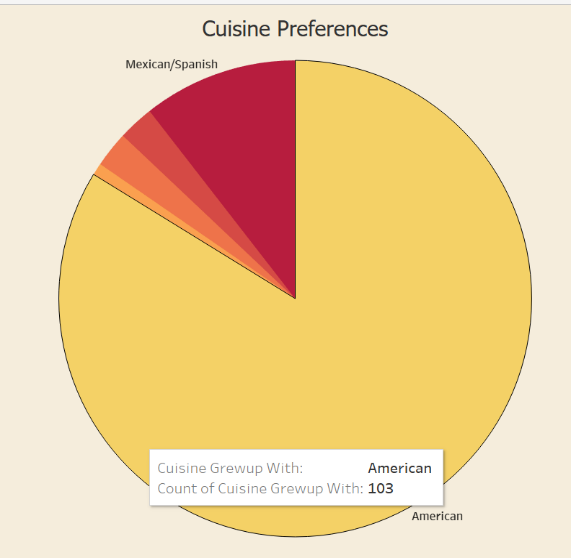
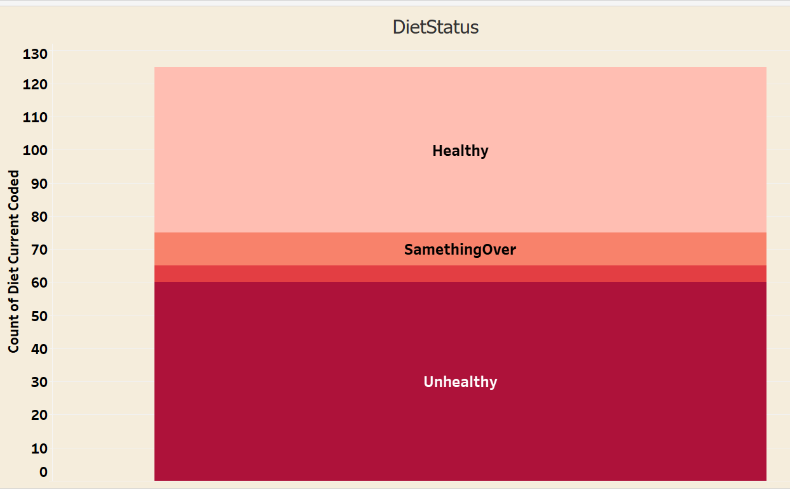
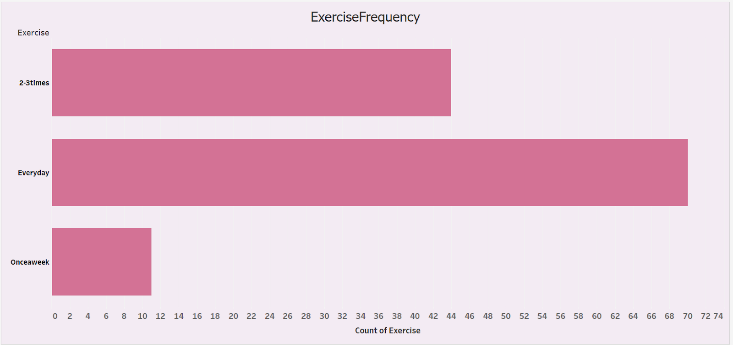
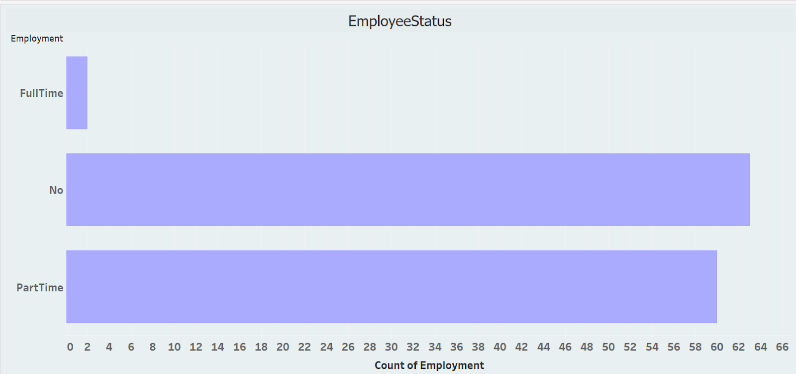
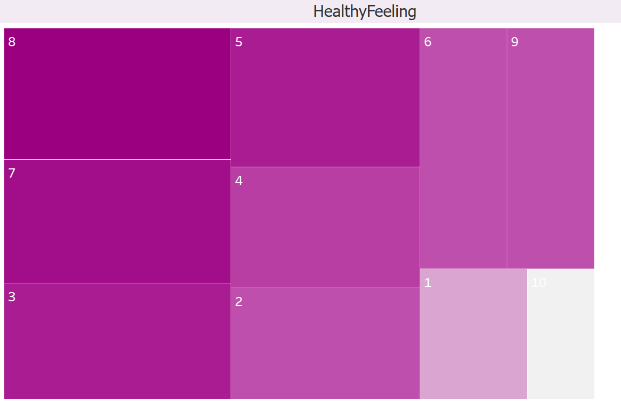
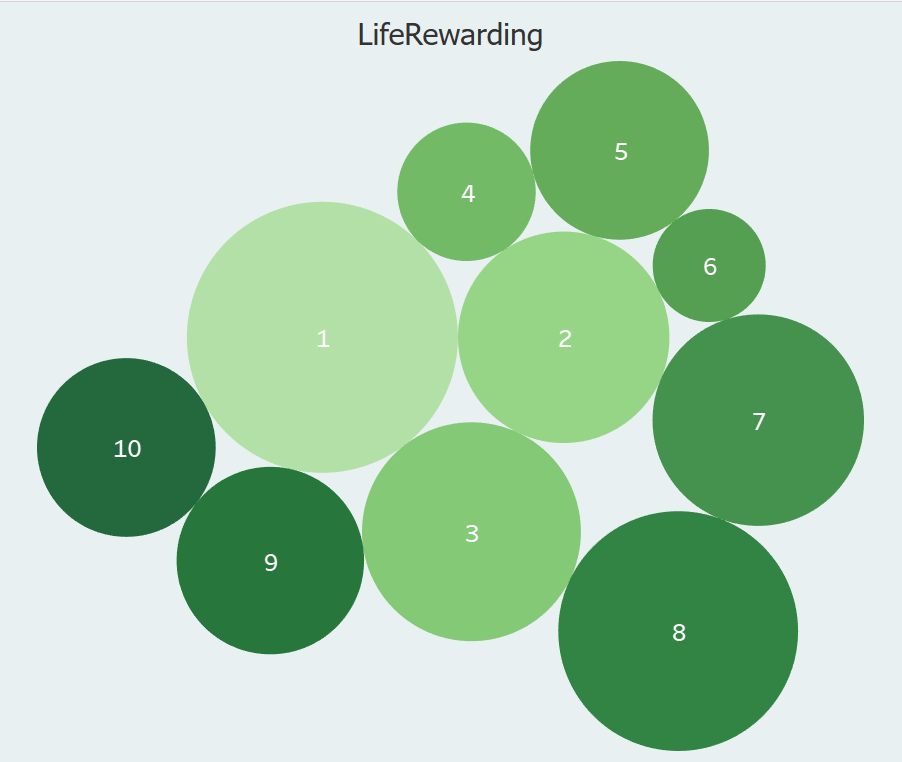
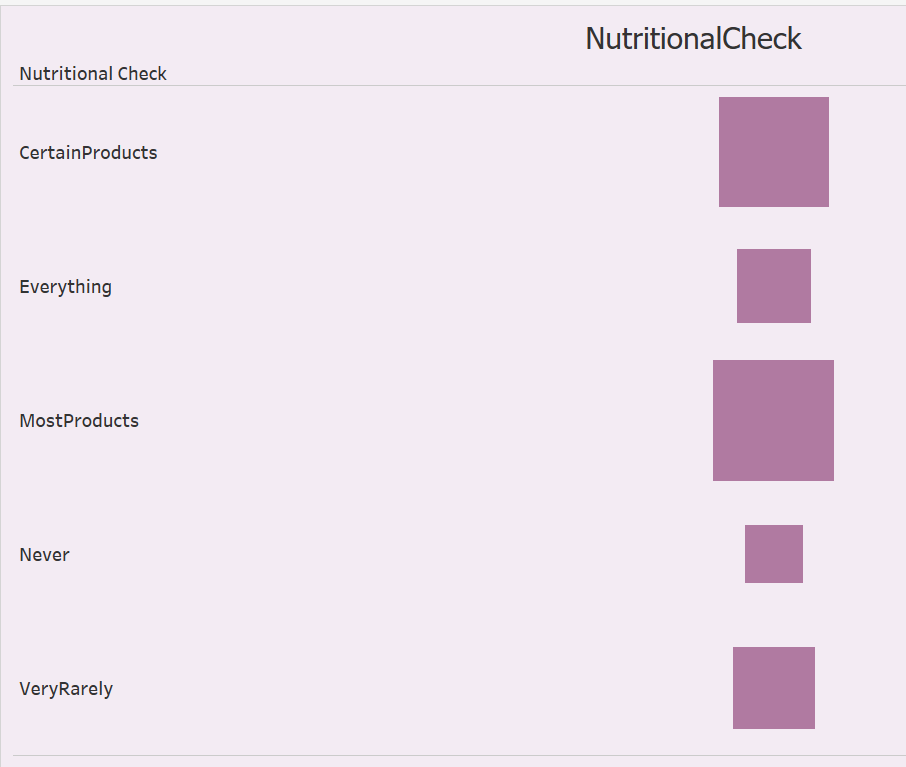
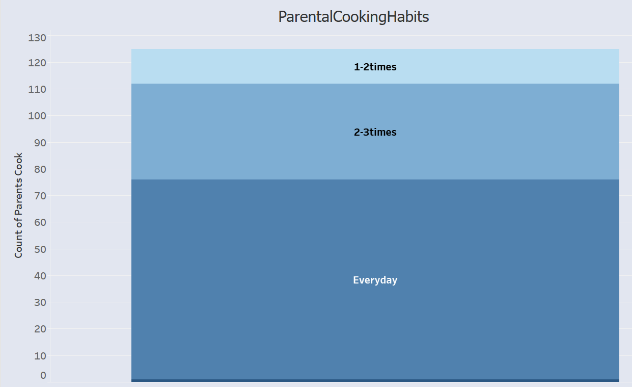
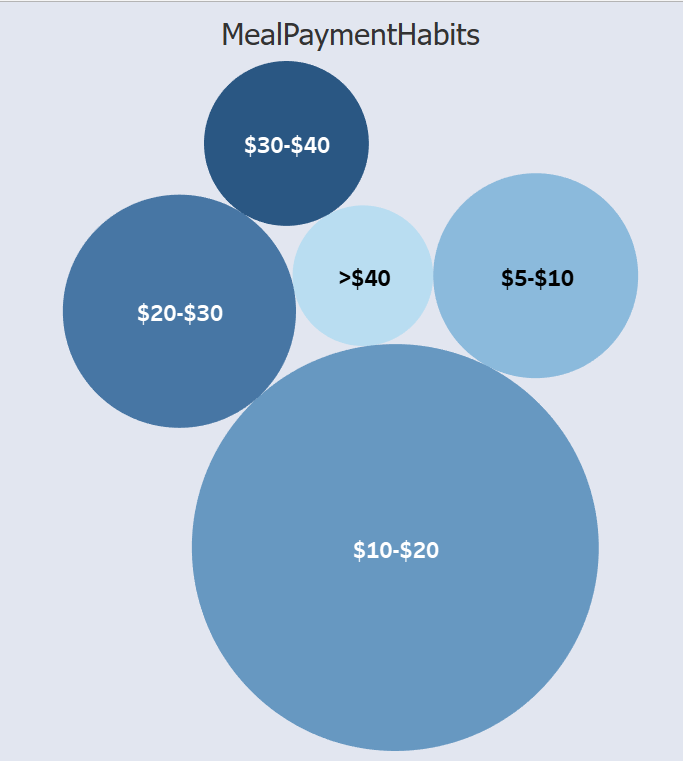
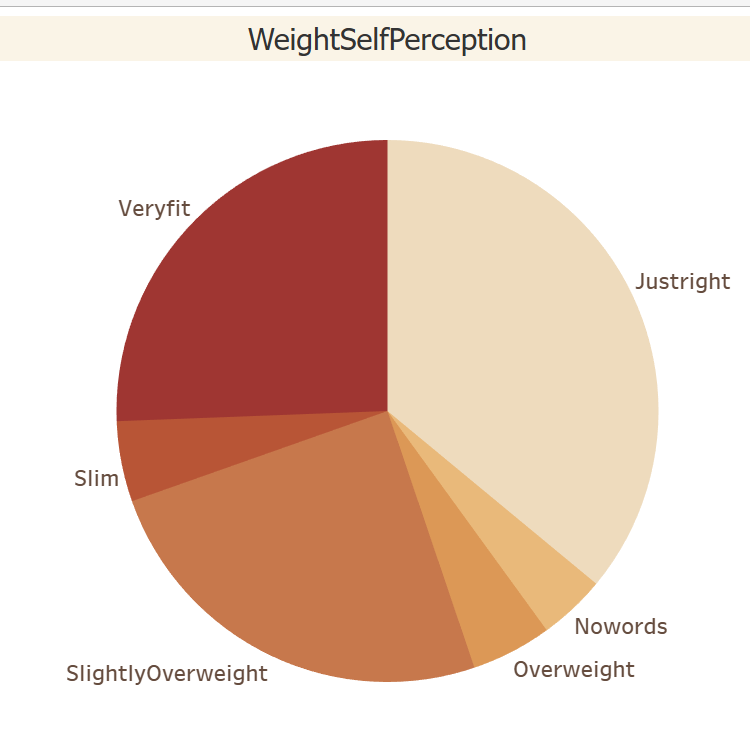
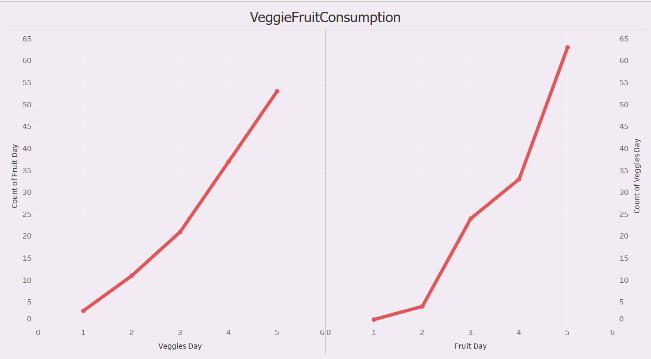
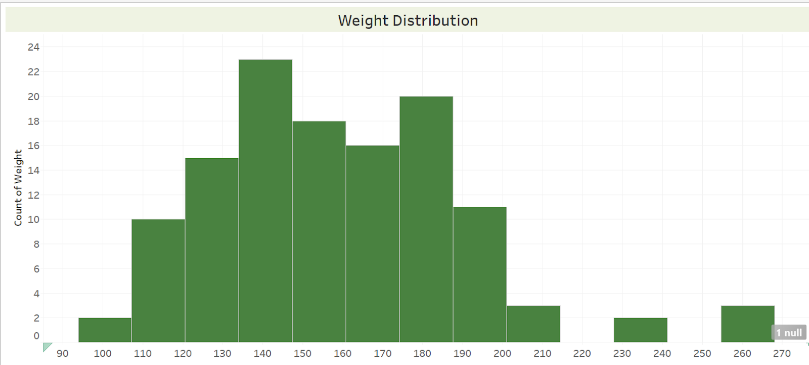
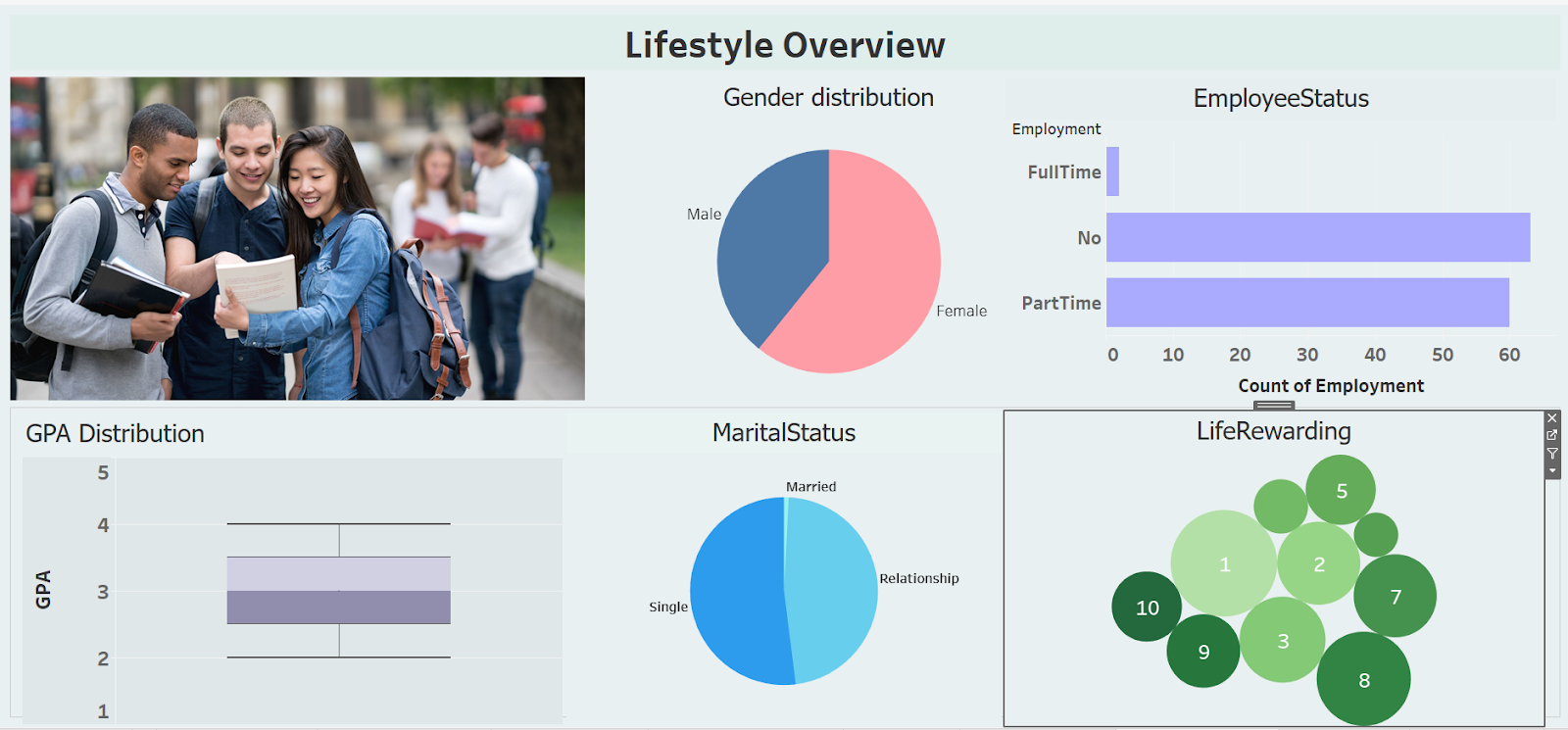
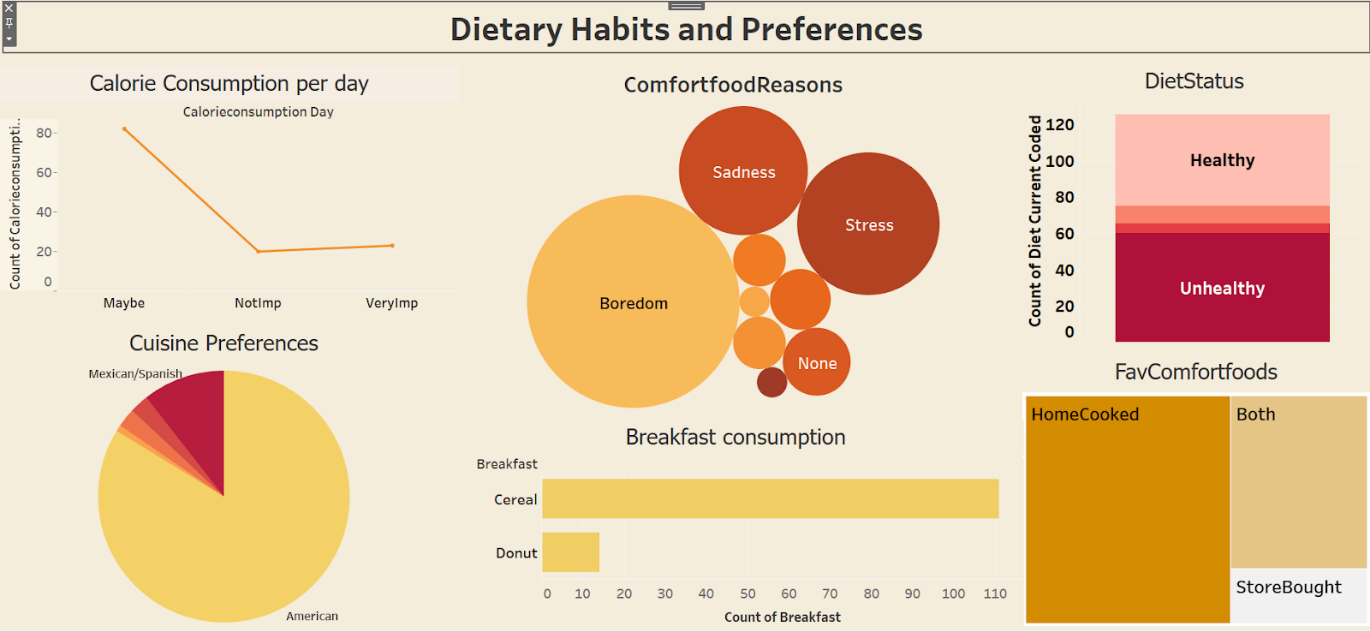
Activity 1.1: Understand the data

Data contains all the meta information regarding the columns described in the CSV files

Column Description of the Dataset:

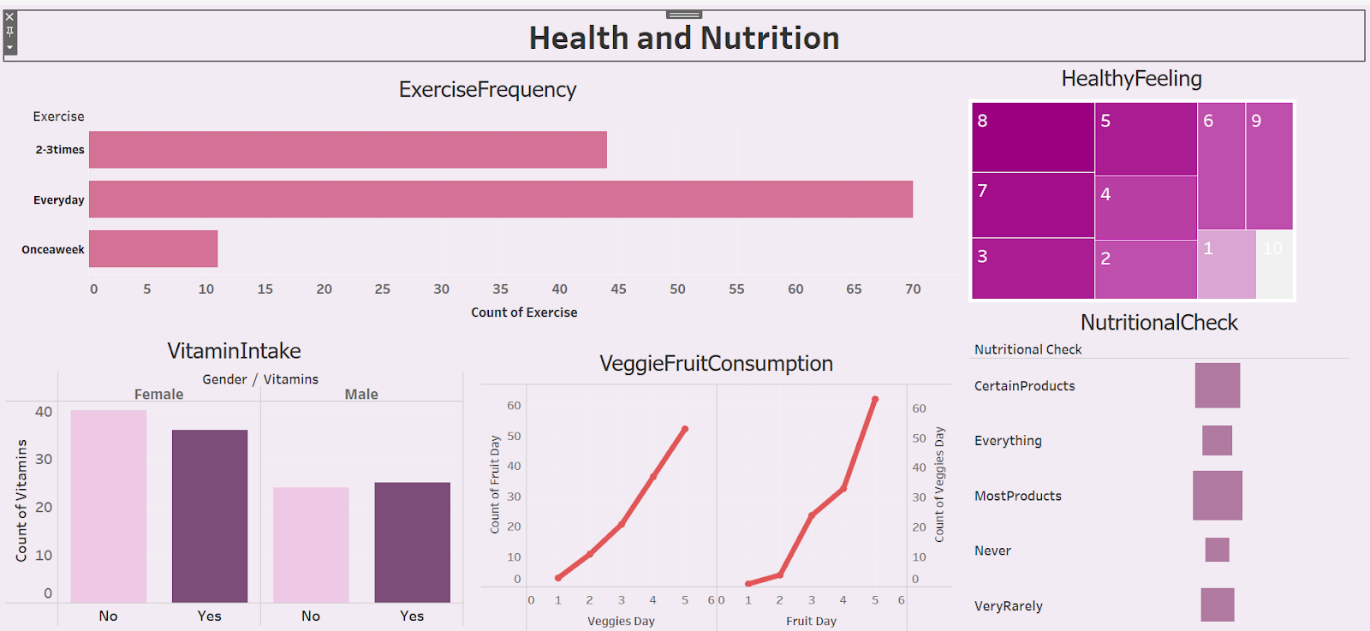
1. GPA :  Grade Point Average of Students
2. Gender : Female, Male
3. Breakfast : Cereal, Donut
4. Calories\_chicken\_piadina : Calories in chicken piadina
5. Calories\_starbuckscone : Calories in starbuckscone
6. Coffee : Espresso, Creamy frapuccino
7. Comfort\_food\_reasons\_coded : reasons behind eating their comfort food
8. Cooking\_per\_week : No.of times students cook per week
9. CuisineGrewUpwith : what type of cuisine they grew up with
10. Diet\_current\_code : type of diet they are following
11. Drink : associated with food, Orange Juice or Soda
12. eatingOutWeekly : No.of times eating out per week
13. Employment : Do students work?
14. Ethnic\_food : how likely do they eat
15. Exercise : how often do students exercise in a week
16. Fav\_cuisine\_coded : current favorite cuisine
17. Fav\_food : homecooked, storebought or both
18. Fruitday : how likely do they eat
19. Grade\_level : which level of schooling they belong to
20. Greekfood : how likely do they eat
21. Healthy\_feeling : “I feel very healthy”- rating, 1- agree,....10-disagree
22. Ideal\_diet\_coded : which diet is ideal for them
23. Income : How much do they earn?
24. Indian\_food : how likely do they eat
25. italian \_food : how likely do they eat
26. Life\_rewarding : “I feel life is rewarding” - rating, 1- agree,....10-disagree
27. Matiral\_status : students marital status
28. Nutritional\_check : how often to they check nutritional values on the food products
29. On\_off\_campus : Living situation - where do they stay
30. Parents\_cook : how often do their parents cook
31. Pay\_meal\_out : how much amount is paid when eaten out
32. Persian\_food : how likely do they eat
33. Self\_perception\_weight : what do they feel about their weight
34. Sports : are they involved in sports
35. Thai\_food :  how likely do they eat
36. Tortilla\_calories : no.of calories in burrito sandwich from chipotle
37. Turkey\_calories : no.of calories in Panera bread roasted and Avacado
38. Veggies\_day : how likely do they eat
39. Vitamins : do they intake vitamins
40. Waffle\_calories : no.of calories in waffles
41. Weight : weight of the student
42. **Data Preparation**
43. Prepare the Data for Visualization
45. Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete. This process helps to make the data easily understandable and ready for creating visualizations to gain insights into the performance and efficiency. Since the data is already cleaned, we can move to visualization.
46. **Explanation video links**
47. Data Loading: https://drive.google.com/file/d/1p6CTgHF-hUzrRe9ZdjfNZr4hgq3a\_8pf/view

**Data Visualization**

* Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.
* **No.of Unique Visualizations**
* Using the given dataset, several unique visualizations can be created to analyze the dietary habits, nutritional intake, and health outcomes of college students. These visualizations include bar charts, line charts, heat maps, scatter plots, pie charts, and maps. These can be used to compare performance, track changes over time, show distribution, identify relationships, breakdown nutritional intake, provide demographic insights, inform resource allocation, and conduct geographical analysis. By leveraging these diverse visualizations, stakeholders can gain comprehensive insights into the dietary choices and health of college students, enabling data-driven decisions to promote better nutrition and overall well-being.
* **visualizations**
* Activity 1.1: Gender Distribution:
* 
* Activity 1.2: GPA Distribution
* 
* Activity 1.3: Breakfast Consumption
* 
* Activity 1.4 : Calorie Consumption per day
* 
* Activity 1.5 :Favorite\_comfort\_foods
* 
* Activity 1.6 :Comfort\_food\_reasons
* 
* Activity 1.7 : Cooking frequency per week
* 
* Activity 1.8 : Cuisine preferences
* 
* Activity 1.9 : Diet Status
* 
* Activity 1.10 :ExerciseFrequency
* 
* Activity 1.11 :Employeestatus
* 
* Activity 1.12 : Healthy Feeling
* 
* Activity 1.13 :LifeRewarding rating
* 
* Activity 1.14 : Marital status
* 
* Activity 1.15 : Nutritional Check
* 
* Activity 1.16 :ParentalCookingHabits
* 
* Activity 1.17 :MealPaymentHabits
* 
* Activity 1.18 :WeightSelfPerception
* 
* Activity 1.19 : Sports participation
* 
* Activity 1.20 :Veggieandfruit\_Consumption
* 
* Activity 1.21 : Vitamin Intake
* 
* Activity 1.22 : Weight
* 
* **Dashboard**
* A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.
* **Responsive and Design of Dashboard**
* Dashboard 1:<https://drive.google.com/file/d/1A-mEgUbPGIWGDCQ3Wn7rA1Wop0Ei6e7w/view?usp=drive_link>
* 
* **Dietary Habits and Preferences**
* Dashboard 2:**Dietary Habits and Preferences**
* Dashboard 2:
* 

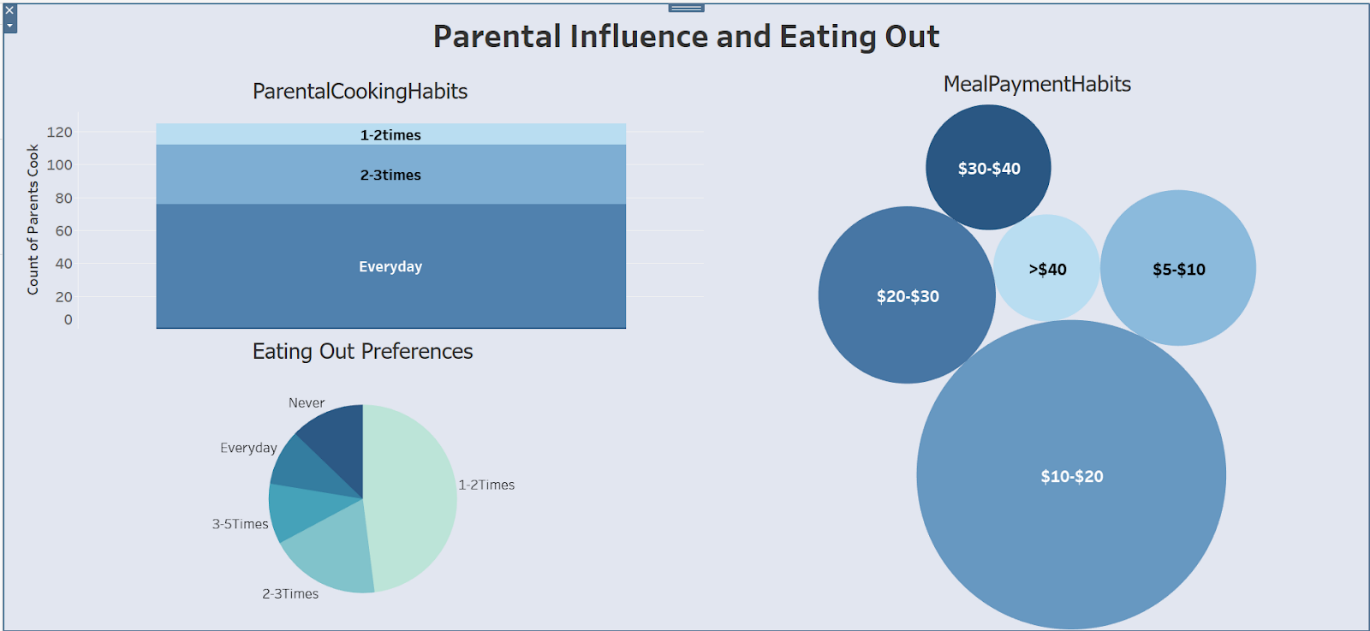
### Health and Nutrition

Dashboard 3: [HealthandNutrition.mp4](https://drive.google.com/file/d/1fZBCXsvgnp0BL12sB8tNY840Pue1orPr/view?usp=drive_link)



### Parental Influence and Eating Out

Dashboard 4:<https://drive.google.com/file/d/1vC8G0WdUG7xxNZbBl75L39-Yg_L_4mSL/view?usp=drive_link>

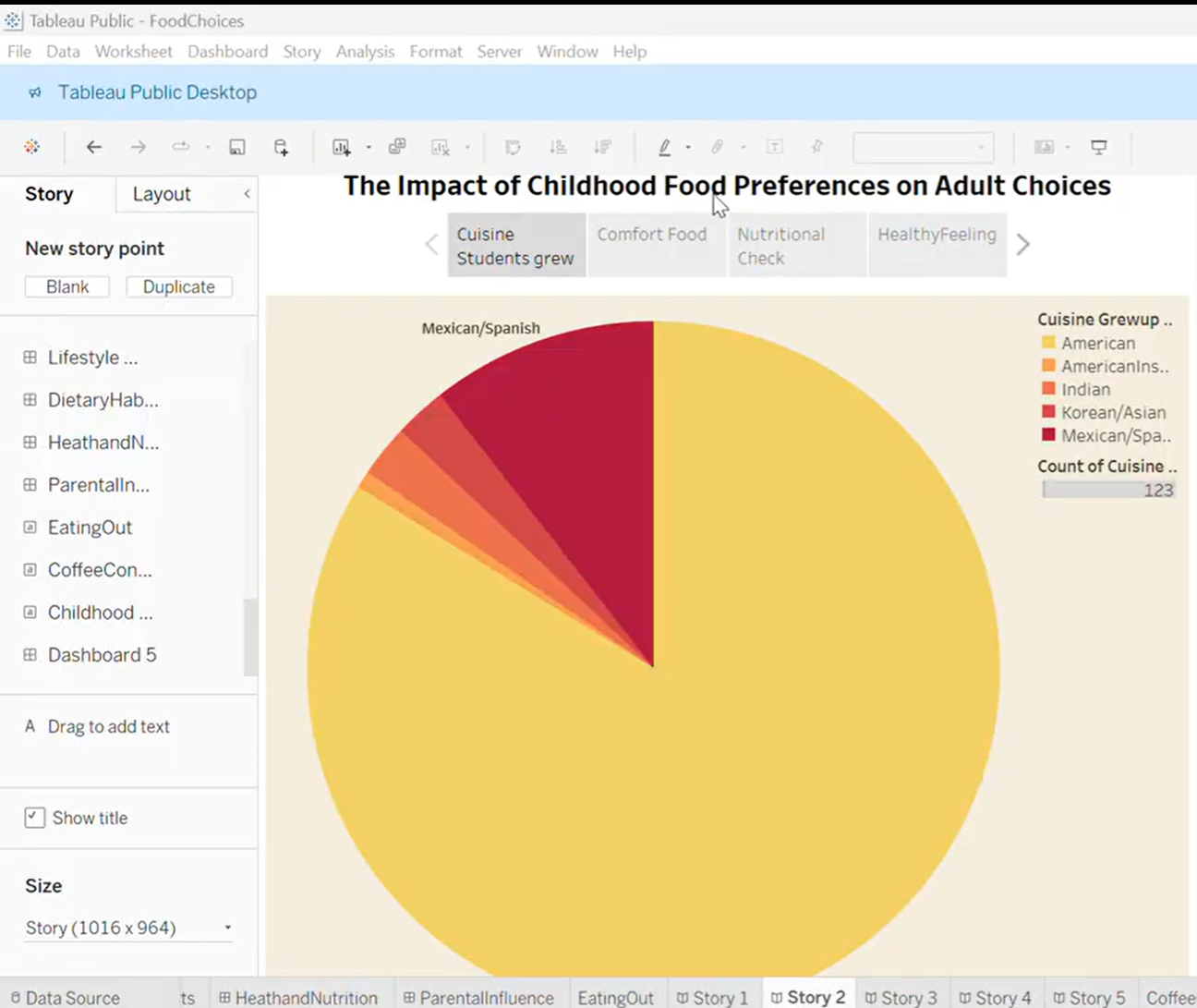


### Story

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

### No of Scenes of Story

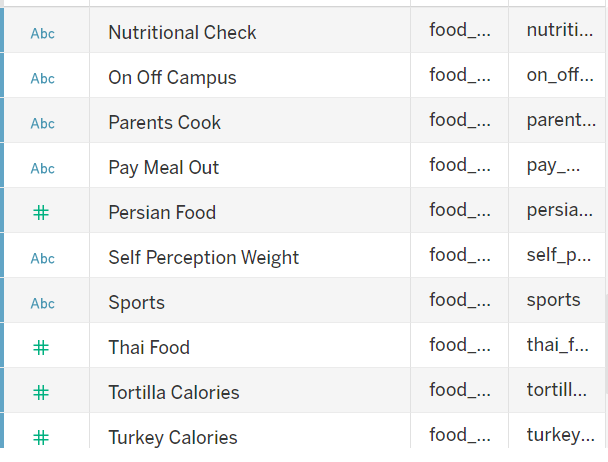
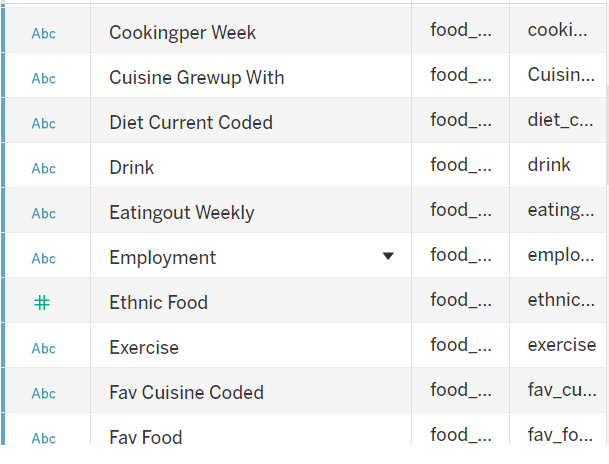
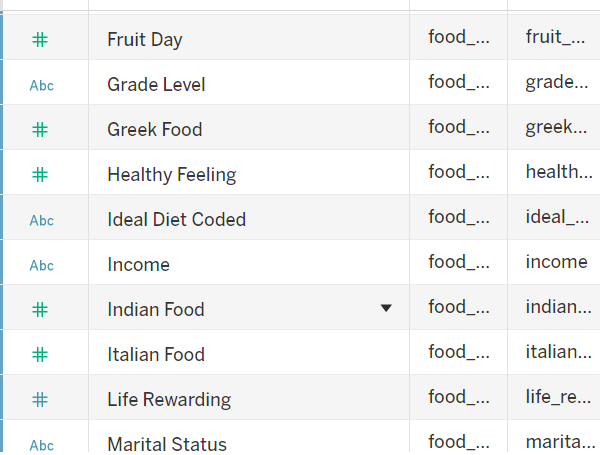
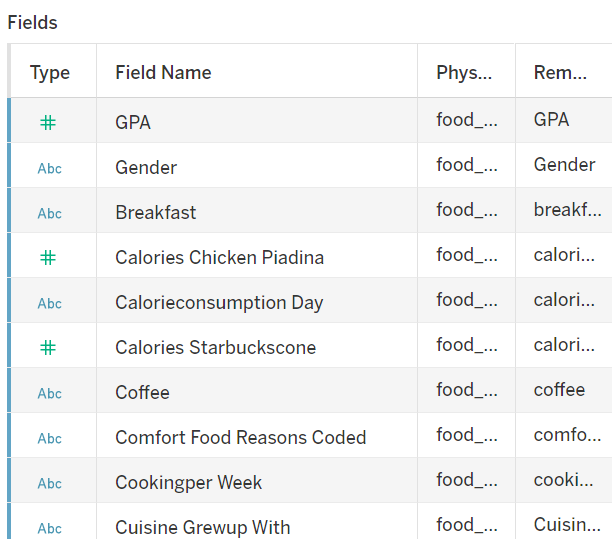
Explanation video link: [Story2.mp4](https://drive.google.com/file/d/1XAYufYK975yShgRiBZNRIiolpspZ6dJ-/view?usp=drive_link)



**Performance Testing**

Amount of Data Loaded

"Amount of Data Loaded" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system.



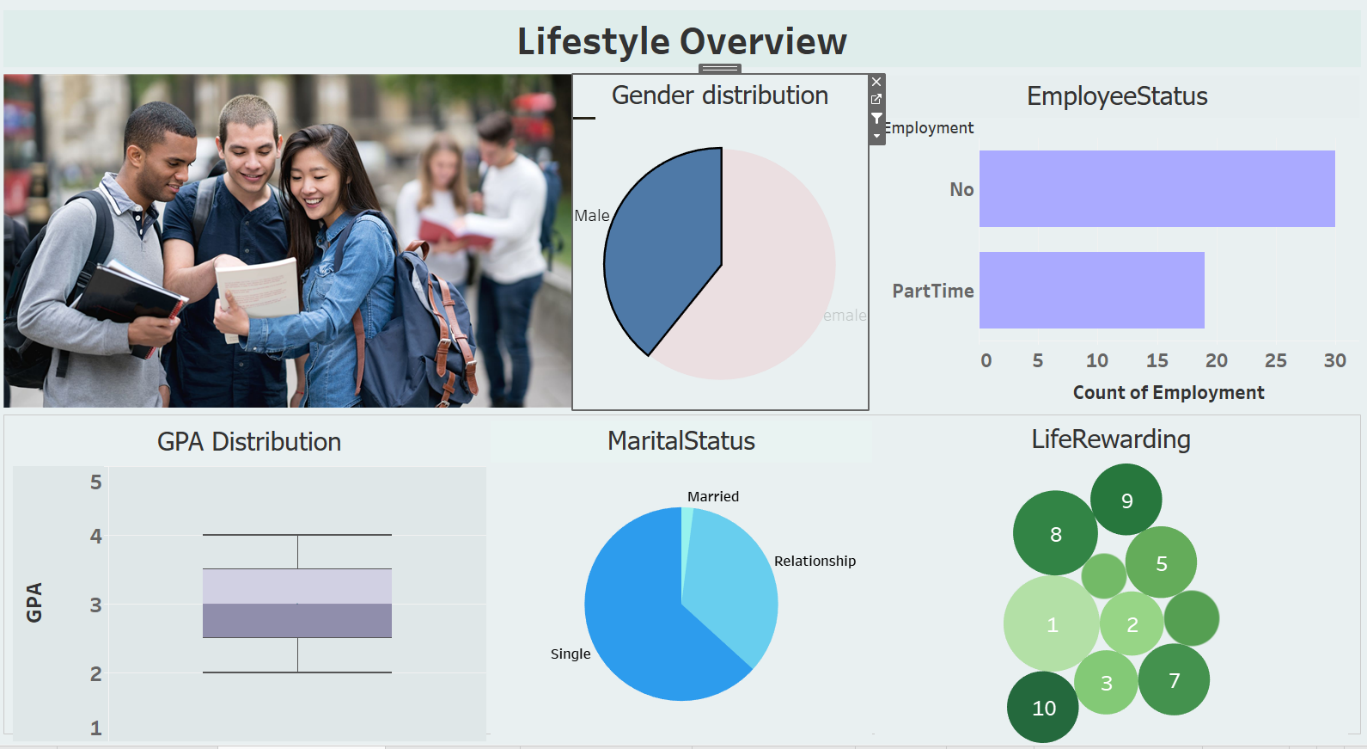
**Utilization of Data Filters**

"Utilization of Filters" refers to the application or use of filters within a system,

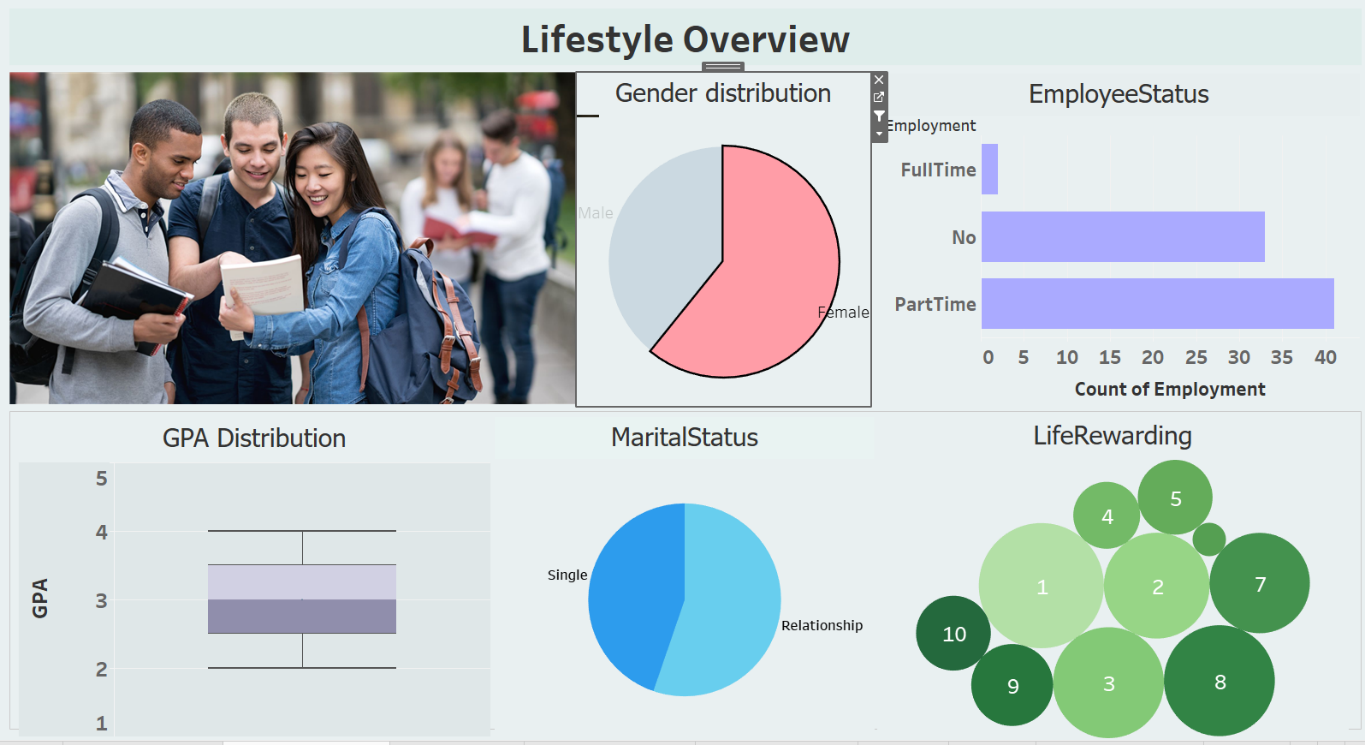
software application, or data processing pipeline to selectively extract, manipulate, or

analyse data based on specified criteria or conditions.

Activity 2.1: Selected “Male” as a Filter



Activity 2.2: Selected “Female” as a Filter



**No of Visualizations/ Graphs**

1. GPA Distribution
2. Gender Distribution
3. Breakfast distribution
4. Calorie Consumption per day
5. Fav Comfort Foods
6. Comfort Food Reasons
7. Cooking Frequency per week
8. Cuisine Preferences
9. Diet Status
10. Exercise Frequency
11. EmployeeStatus
12. HealthyFeeling
13. LifeRewardingRating
14. Marital Status
15. Nutritional Check
16. ParentalCookingHabits
17. MealPaymentHabits
18. WeightSelfPerception
19. SportsParticipation
20. VitaminIntake
21. WeightDistribution
22. Eatingout
23. Coffee Consumption

### Web integration

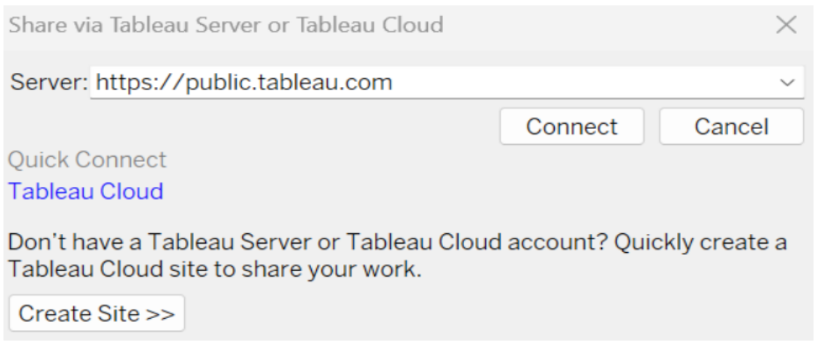
Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others. Publishing dashboard and reports to tableau public

**Go to Dashboard/story, click on share button on the top ribbon**

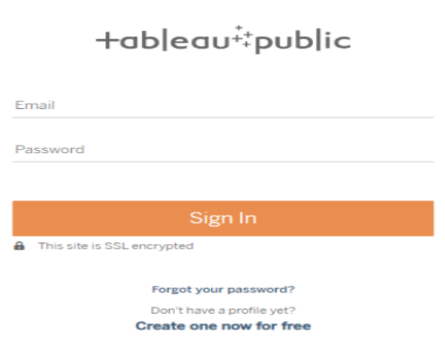
Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others. Publishing dashboard and reports to tableau public

Step 1: Go to Dashboard/story, click on share button on the top ribbon

Give the server address of your tableau public account and click on connect.



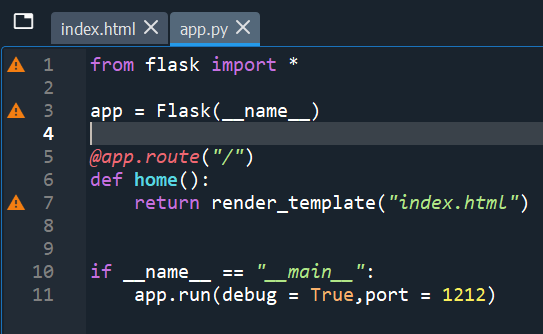
Step 2: Once you click on connect it will ask you for tableau public user name and password.



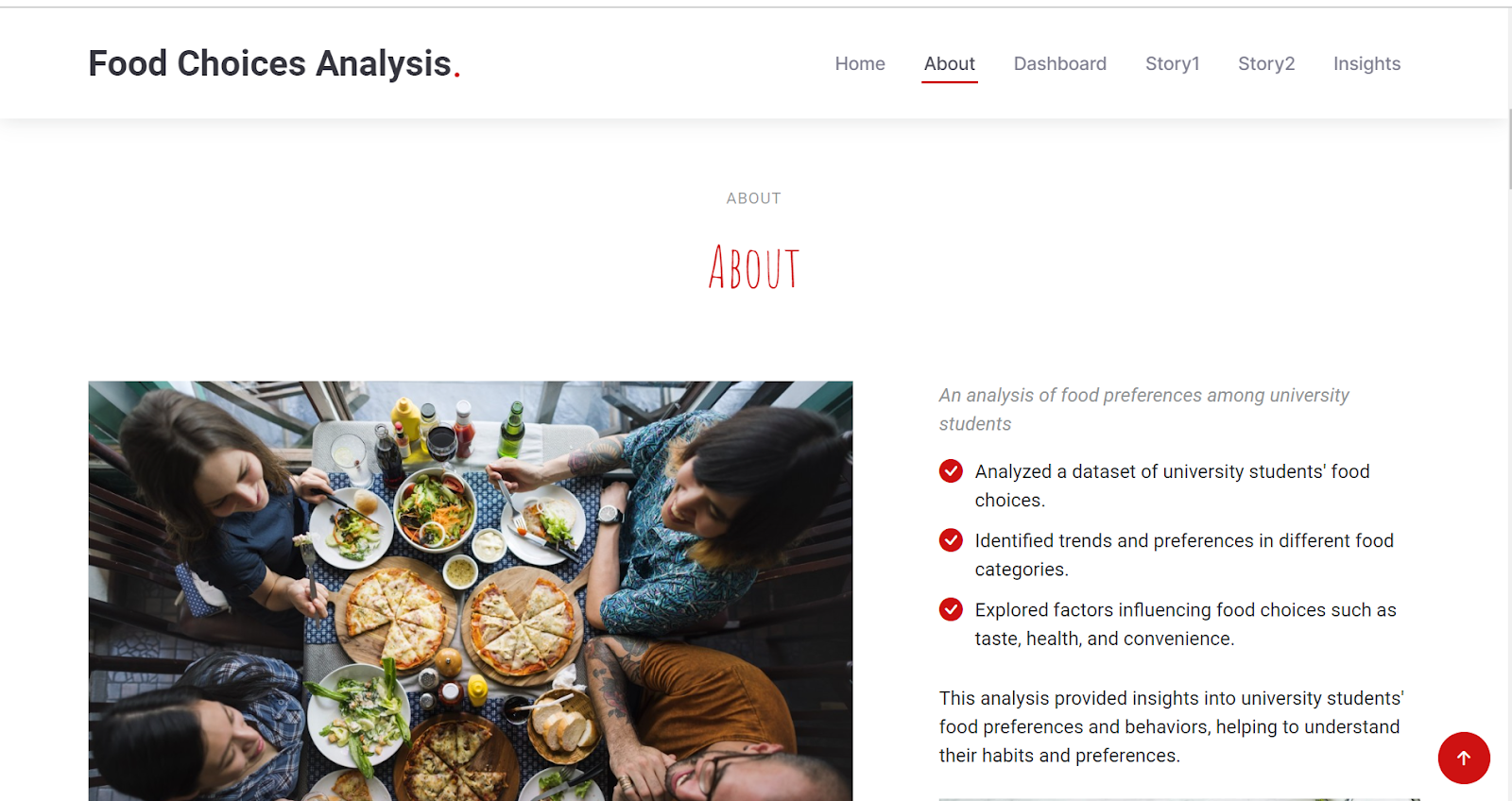
Once you login into your tableau public using the credentials, the particular visualization will be published into tableau public.

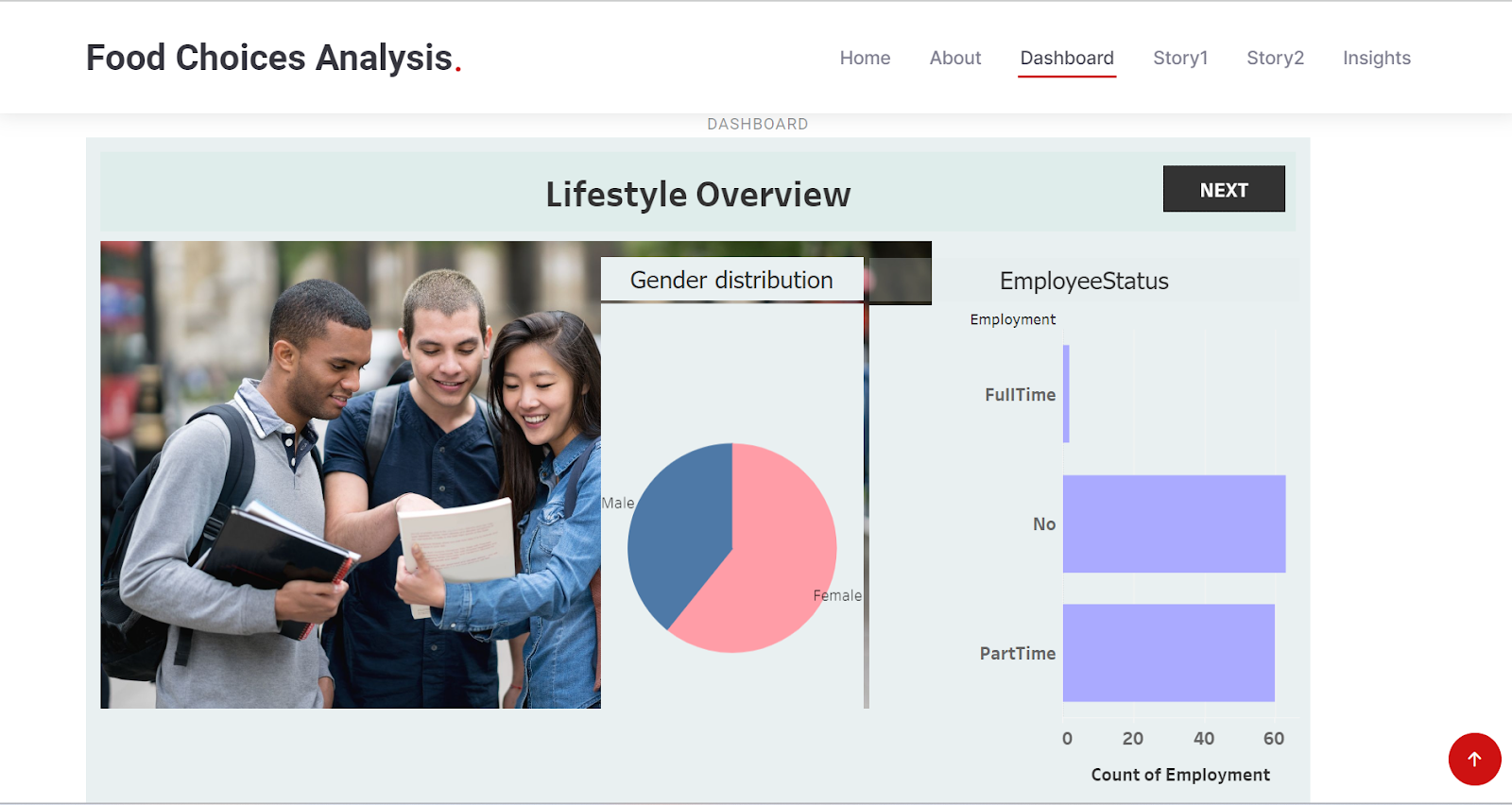
Note: While publishing the visualization to the public, the respective sheet will get published when you click on share option.

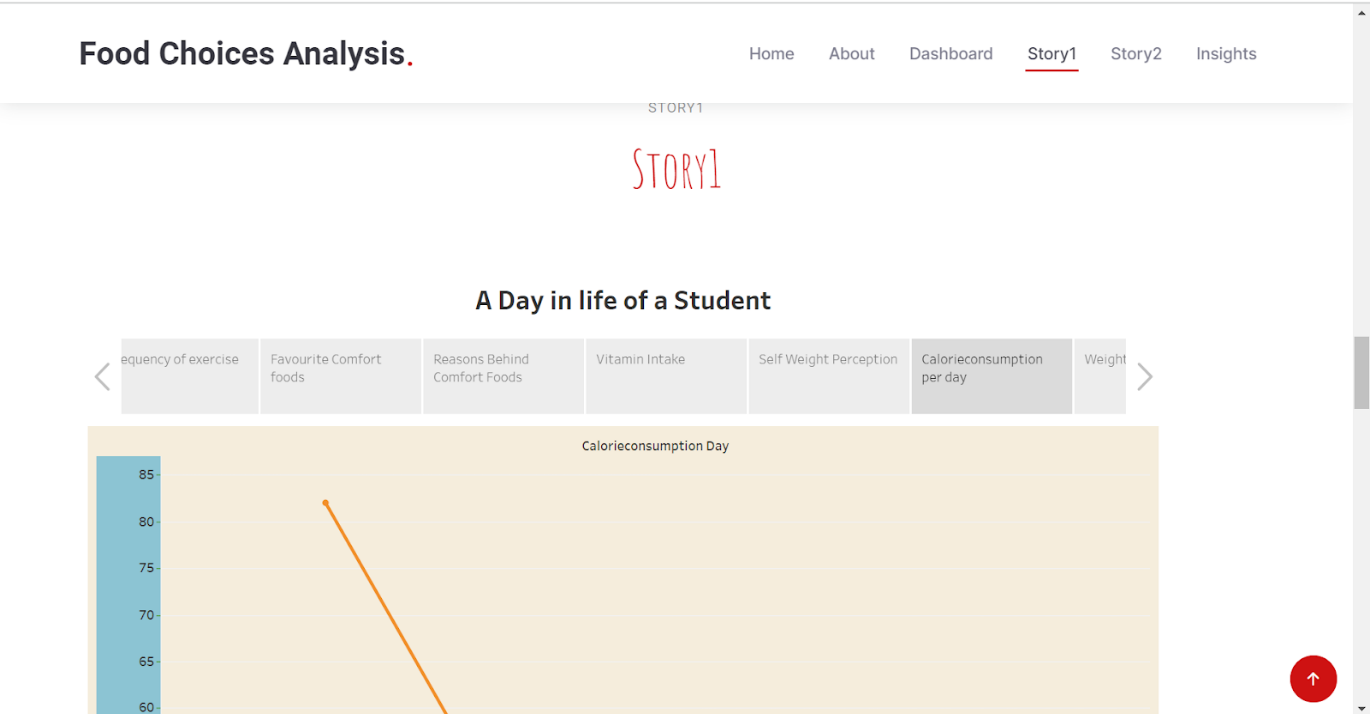
**Dashboard and Story embed with UI With Flask**



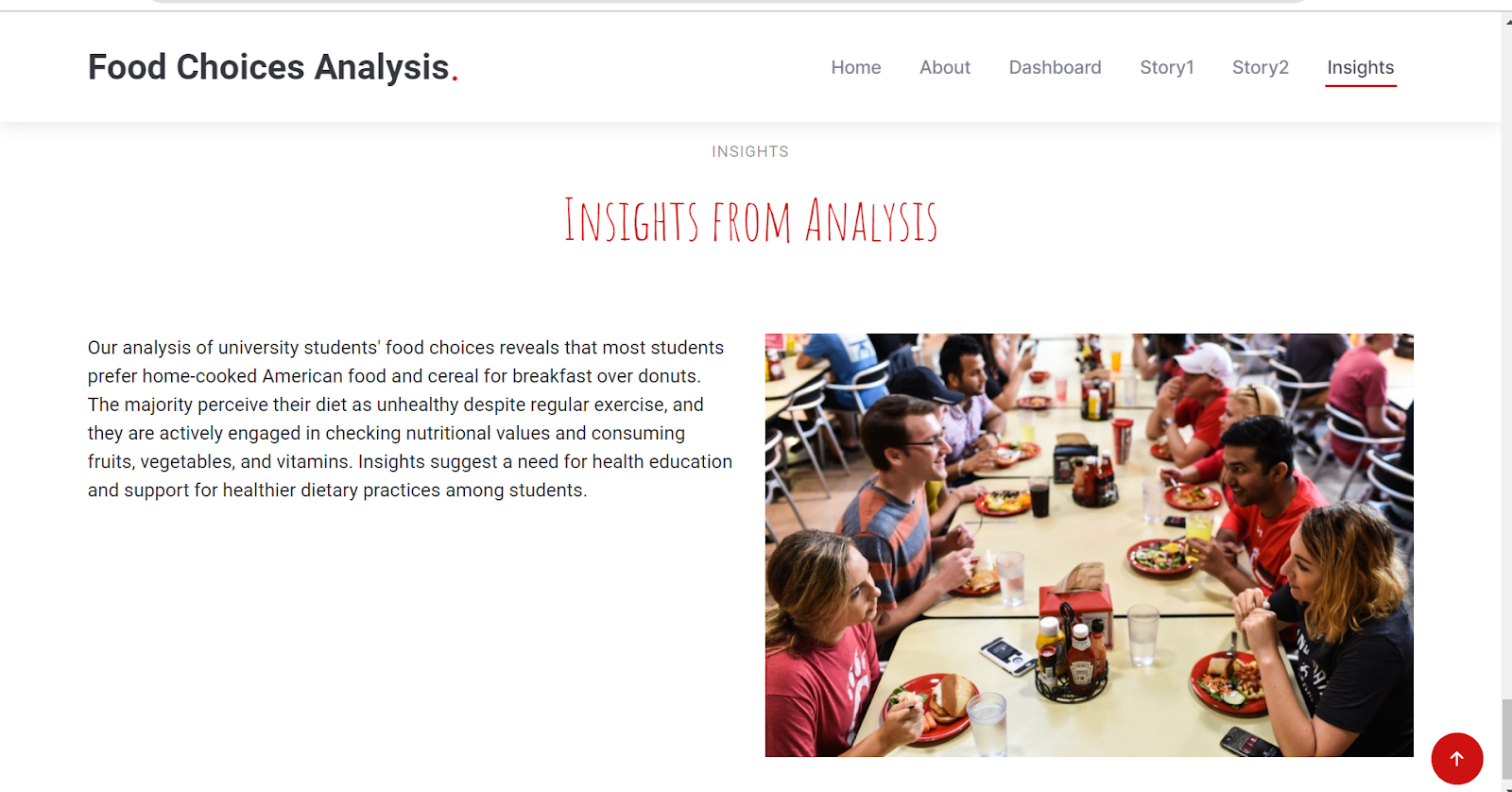


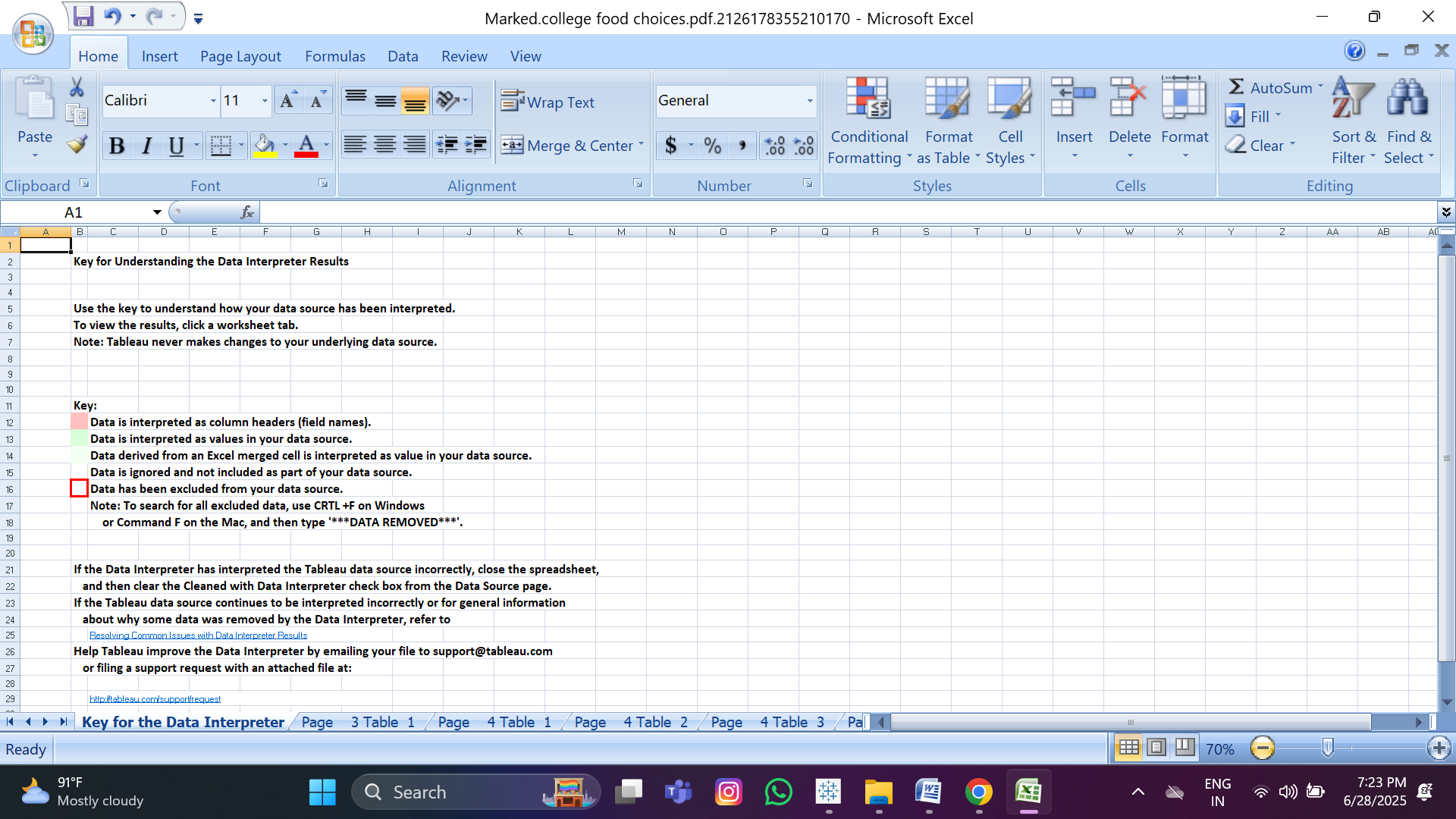










* 

Clean Data from Excel, CSV, PDF, and Google Sheets with Data Interpreter

*Applies to: Tableau Cloud, Tableau Desktop, Tableau Server*

When you track data in Excel spreadsheets, you create them with the human interface in mind. To make your spreadsheets easy to read, you might include things like titles, stacked headers, notes, maybe empty rows and columns to add white space, and you probably have multiple tabs of data too.

When you want to analyze this data in Tableau, these aesthetically pleasing attributes make it very difficult for Tableau to interpret your data. That’s where Data Interpreter can help.

**Tip:** Though Tableau's Excel add-in is no longer supported, Data Interpreter can help you reshape your data for analysis in Tableau.

What does Data Interpreter do?

Data Interpreter can give you a head start when cleaning your data. It can detect things like titles, notes, footers, empty cells, and so on and bypass them to identify the actual fields and values in your data set.

It can even detect additional tables and sub-tables so that you can work with a subset of your data independently of the other data.

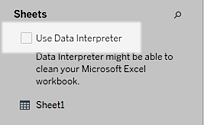
After Data Interpreter has done its magic, you can check its work to make sure it captured the data that you wanted and identified it correctly. Then, you can make any necessary adjustments.

After you select the data that you want to work with, you might also need to do some additional cleaning steps like pivoting your data, splitting fields, or adding filters to get the data in the shape you want before starting your analysis.

**Note**: If your data needs more cleaning than what Data Interpreter can help you with, try [Tableau Prep(Link opens in a new window)](https://www.tableau.com/products/prep).

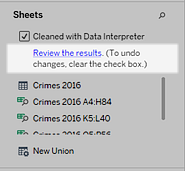
Turn on Data Interpreter and review results

1. From the **Connect** pane, connect to an Excel spreadsheet or other connector that supports Data Interpreter such as Text (.csv) files, PDF files or Google sheets.
2. Drag a table to the canvas (if needed), then on the **Data Source** page, in the left pane, select the **Use Data Interpreter** check box to see if Data Interpreter can help clean up your data.

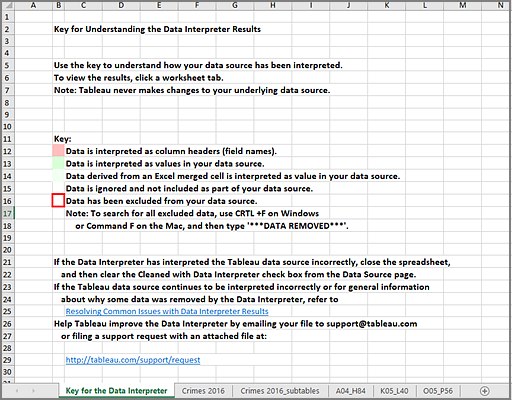


**Note:** When you clean your data with Data Interpreter, Data Interpreter cleans all the data associated with a connection in the data source. Data Interpreter does not change the underlying data.

1. In the Data pane, click the **Review the results** link to review the results of the Data Interpreter.



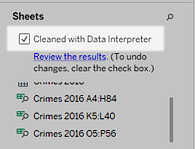
A copy of your data source opens in Excel on the **Key for the Data Interpreter** tab. Review the key to find out how to read the results.



1. Click each tab to review how Data Interpreter interpreted the data source.

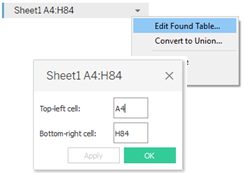
If Data Interpreter found additional tables, also called found tables or sub-tables, they are identified in the <sheet name>\_subtables tab by outlining their cell ranges. A separate tab is also included for each sub-table, color coded to identify the header and data rows.

If Data Interpreter does not provide the expected results, clear the **Cleaned with Data Interpreter** check box to use the original data source.



1. To replace the current table with any of the found tables, drag the current table off the canvas and then drag the found table that you want to use to the canvas.

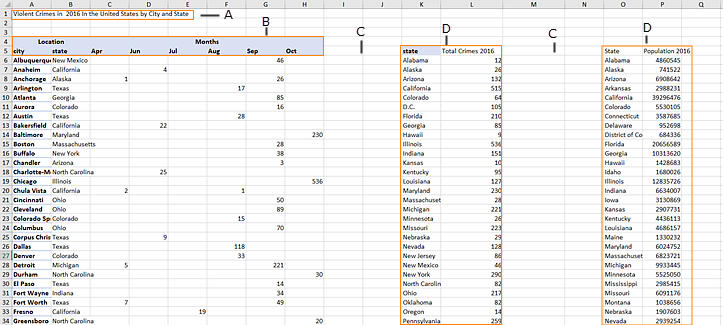
If Data interpreter has misidentified the range of the found table, after you drag the found table to the canvas, click the drop-down arrow on that table, and then select **Edit Found Table** to adjust the corners of the found table (the top-left cell and bottom-right cell of the table).



1. After you have the data that you want to work with, you can apply any additional cleaning operations to your data so that you can analyze it.

Data Interpreter Example

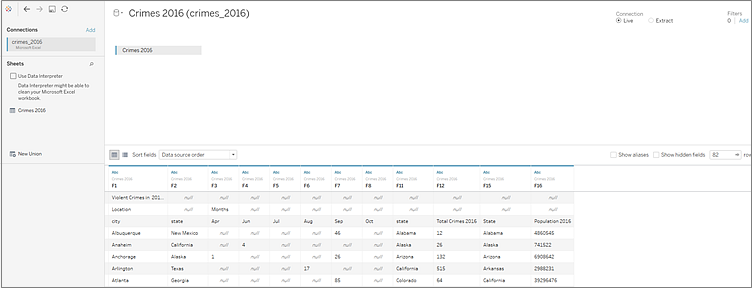
In this example we are connecting to an Excel spreadsheet with violent crime data by city and state for the year 2016. This spreadsheet includes multiple tables on one sheet and some extra formatting.



1. Title
2. Merged header cells
3. Extra white space
4. Sub-tables

The extra formatting in this spreadsheet makes it difficult for Tableau to determine what the field headers and values are.

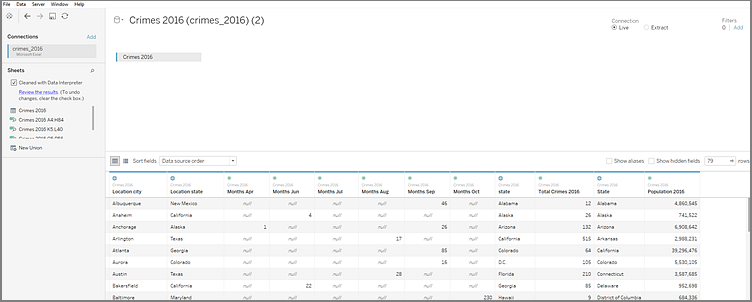
Instead, it reads the data vertically and assigns each column the default value F1, F2, F3 (Field 1, Field 2, Field 3) and so on. Blank cells are read as null values.



To see if Data Interpreter can help clean this data set, we select **Use Data Interpreter**.

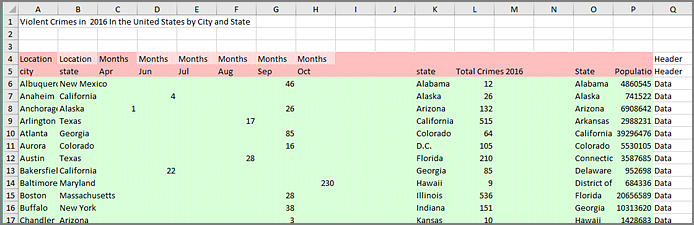
Data Interpreter detected the proper headings for the fields, removed the extra formatting and found several sub-tables. The sub-tables are listed in the **Sheets** section in the Data pane and are named using the original sheet name and the cell ranges for each sub-table.

In this example there are three sub-tables: **Crimes 2016 A4:H84**, **Crimes 2016 K5:L40**, and **Crimes 2016 O5:P56**.

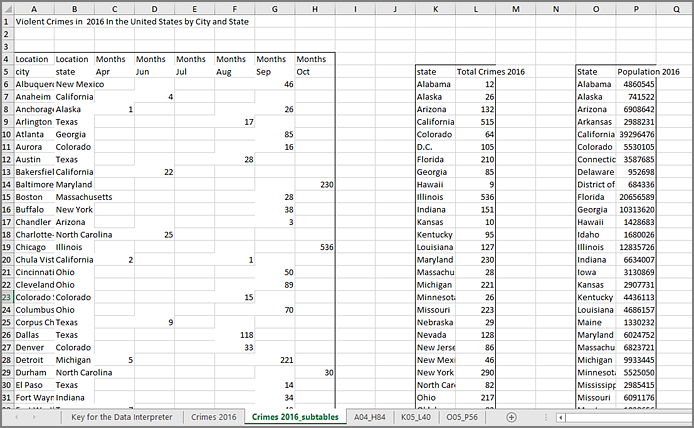


To examine the results of the Data Interpreter more closely, we click the **Review the results** link in the Data pane to view an annotated copy of the spreadsheet.

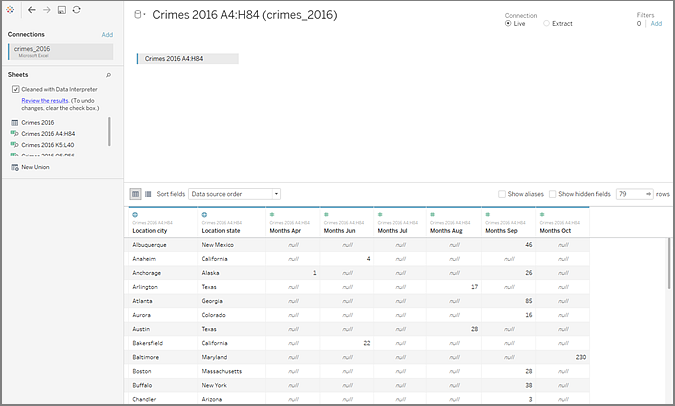
Here we see a copy of the original data, color coded to identify which data was identified as header data and which data was identified as field values.



The next tab shows us the sub-tables that Data Interpreter found, outlined by the cell ranges.



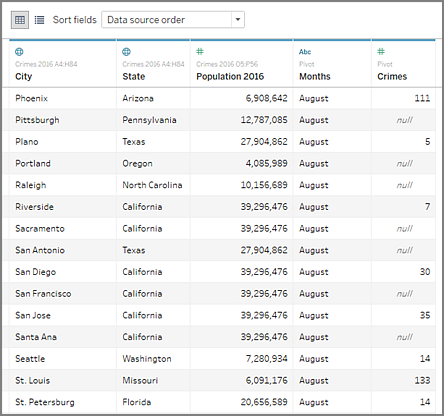
In this example the first sub-table, **Crimes 2016 A4:H84**, has the main data that we want to work with. To use this table as our data table, we can simply drag the original table off the canvas and then drag the new table to the canvas.



Once we have the data that we want to work with in the canvas, we can do some additional clean up on the data. For example we can:

* Change the field names so that they represent city, state, and month names.
* Pivot the months fields.
* Drag in the third sub-table **Crimes 2016 o5:P56** and join it to our first sub-table on the **State** field to include state populations for our analysis.
* Hide any duplicate fields that were added as a result of the join.

The results might look something like this:



Now we are ready to start analyzing our data in Tableau.

When Data Interpreter is not available

The Data Interpreter option might not be available for the following reasons:

* **The data source is already in a format that Tableau can interpret:** If Tableau Desktop doesn't need extra help from Data Interpreter to handle unique formatting or extraneous information, the Data Interpreter option is not available.
* **Many rows or many columns:** The Data Interpreter option is not be available when your data has the following attributes:
  + Data contains more than 2000 columns.
  + Data contains more than 3000 rows and more than 150 columns.
* **The data source is not supported:** Data Interpreter is only available for Microsoft Excel, Text (.csv) files, PDF files and Google Sheets. For Excel, your data must be in the .xls or .xlsx format.