## Exploratory Data Analysis in Python

Exploratory data analysis (EDA) is a critical initial step in the data science workflow. It involves using Python libraries to inspect, summarize, and visualize data to uncover trends, patterns, and relationships. Here’s a breakdown of the key steps in performing EDA with Python:

1. Importing Libraries:

* pandas (pd): For data manipulation and analysis.
* NumPy (np): For numerical computations.
* Matplotlib.pyplot (plt): For basic plotting functionalities.
* Seaborn (sns): A built-on top of Matplotlib, providing high-level visualization.

2. Loading the Data:

* Use pd.read\_csv() for CSV files, similar functions exist for other data formats (e.g., .xlsx, .json).

3. Initial Inspection:

* Get an overview of the data using df.head(), .tail(), and .info().
* Check data types with df.dtypes.

4. Data Cleaning:

* Identify and handle missing values using methods like df.isnull().sum().
* Find and address duplicates with df.duplicated().sum().

5. Univariate Analysis:

* Analyze single variables at a time.
* Use descriptive statistics with df.describe() for numerical data.
* Create histograms, box plots, and density plots to visualize distributions.

6. Bivariate Analysis:

* Explore relationships between two variables.
* Create scatter plots to identify trends and potential correlations.

7. Visualization:

* Effective visualizations are crucial for understanding data.
* Use various plots like bar charts, pie charts, and heatmaps to represent categorical data.

## Conclusion

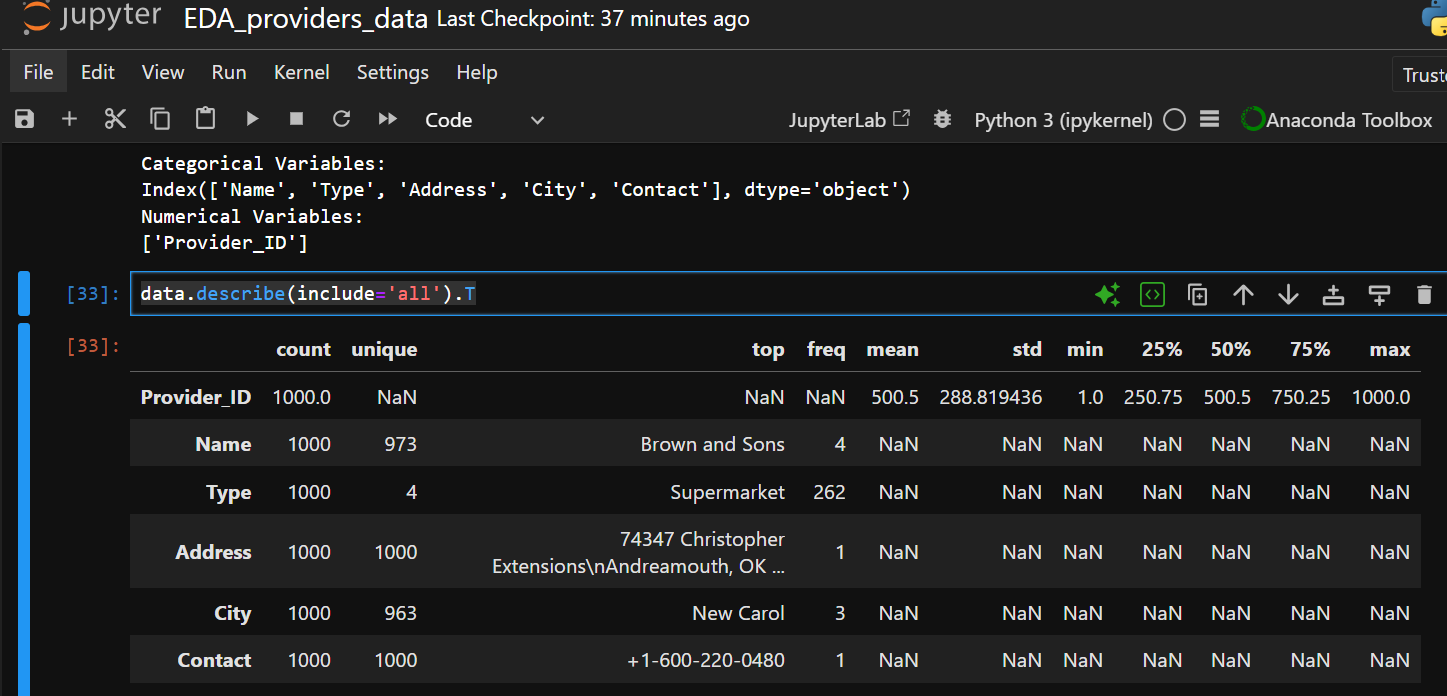
Exploratory Data Analysis (EDA) is crucial for understanding datasets, identifying patterns, and informing subsequent analysis. Data pre-processing and feature engineering are essential steps in preparing data for analysis, involving tasks such as data reduction, cleaning, and transformation. [Python libraries](https://www.analyticsvidhya.com/blog/2024/04/python-libraries/) offer powerful tools for executing these steps efficiently.Also, in the article we talk about how eda using python and you can make to it we showed a complete guide for that.

we tried to analyze the factors involved with food providers, receivers and claims.

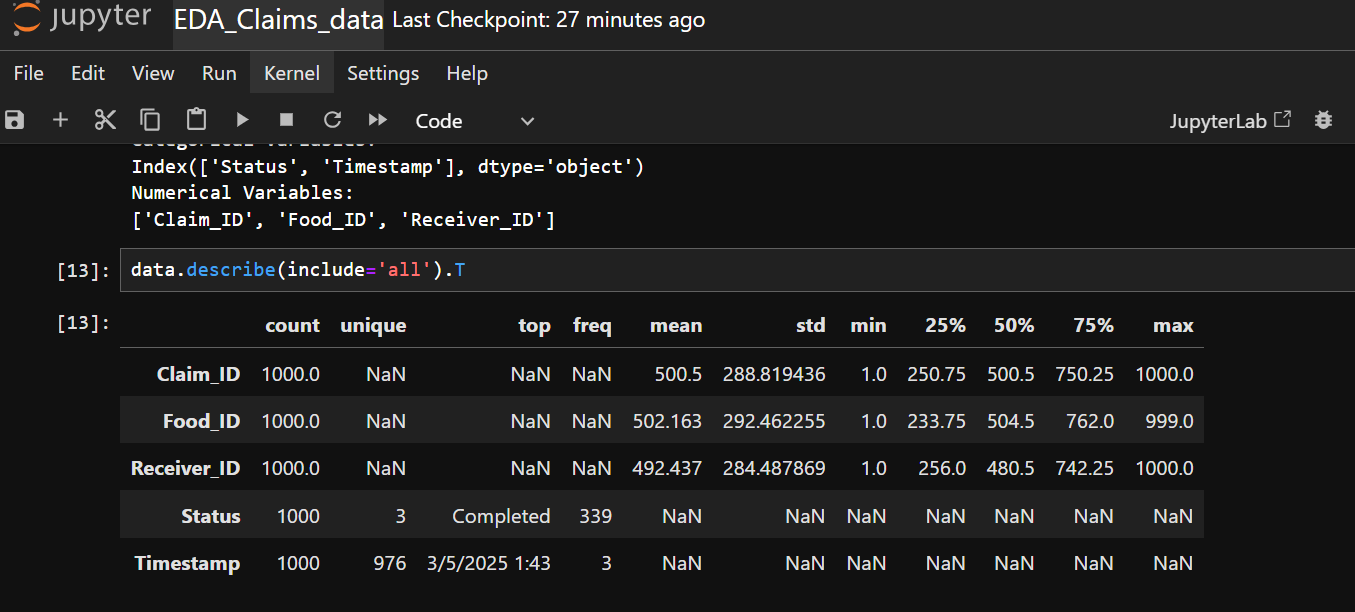
* Data Analysis helps to find the basic structure of the dataset.
* Performed Feature Engineering by adding some columns which contribute to our analysis.
* Data transformations normalize the columns.
* We used different visualizations for Exploratory data analysis (EDA) like Univariate, Bi-Variate, and Multivariate Analysis.

Through EDA, we got useful insights, and below are the factors helps to reduce food wastage and a few takeaways:

1. Supermarkets are listing more surplus foods and Brown and Sons are the top providers and contributed 4 times.



2) Totally 339 people claim the food and status is completed.

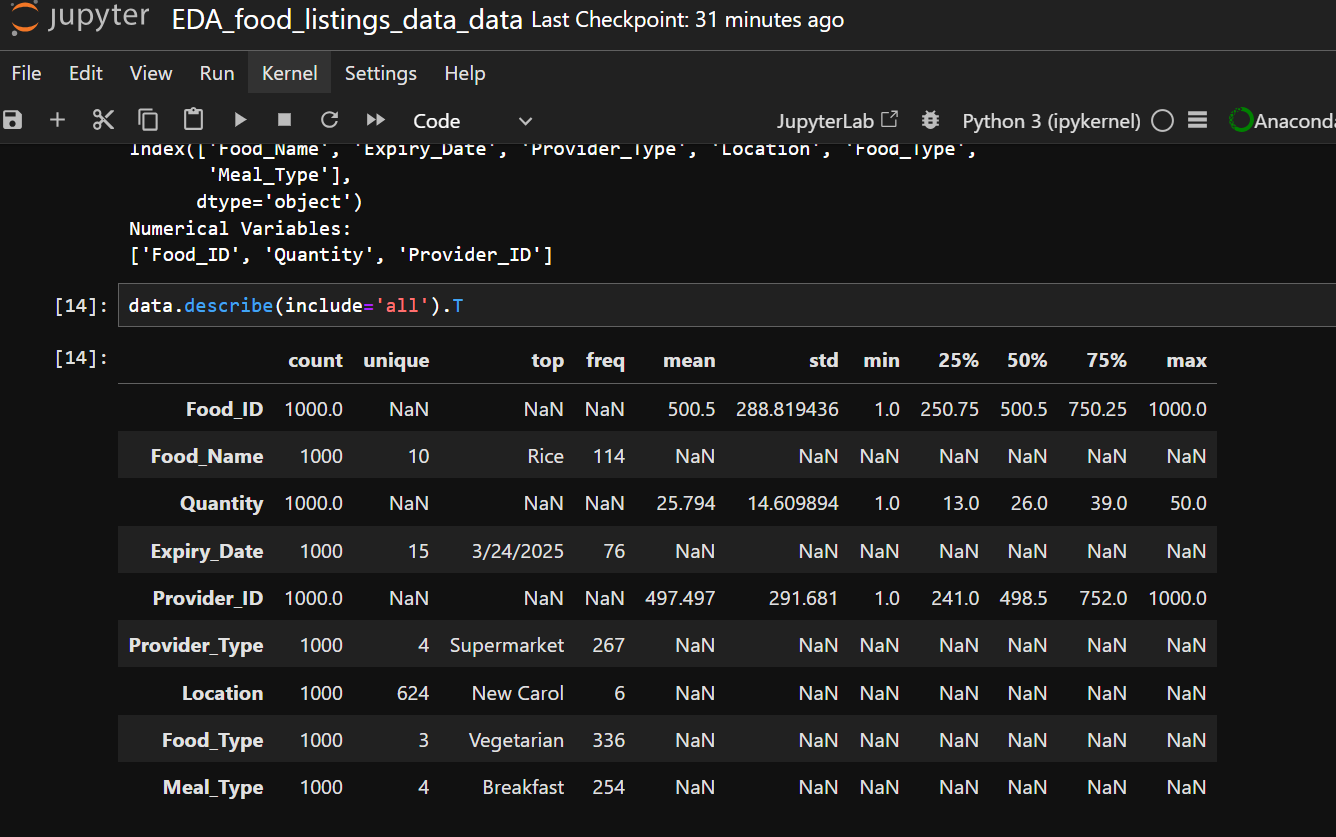


3) Rice is the top contributed food.

The most listed food type is Veg.

The most consumed food is “breakfast.”

New carol is the top location listed food listing for 6 times than any other location.



4)

Heather Brown is the top receiver and received two times.

The top type is NGO and the frequency is 274.

The receivers are from New christopher for 3 times.

