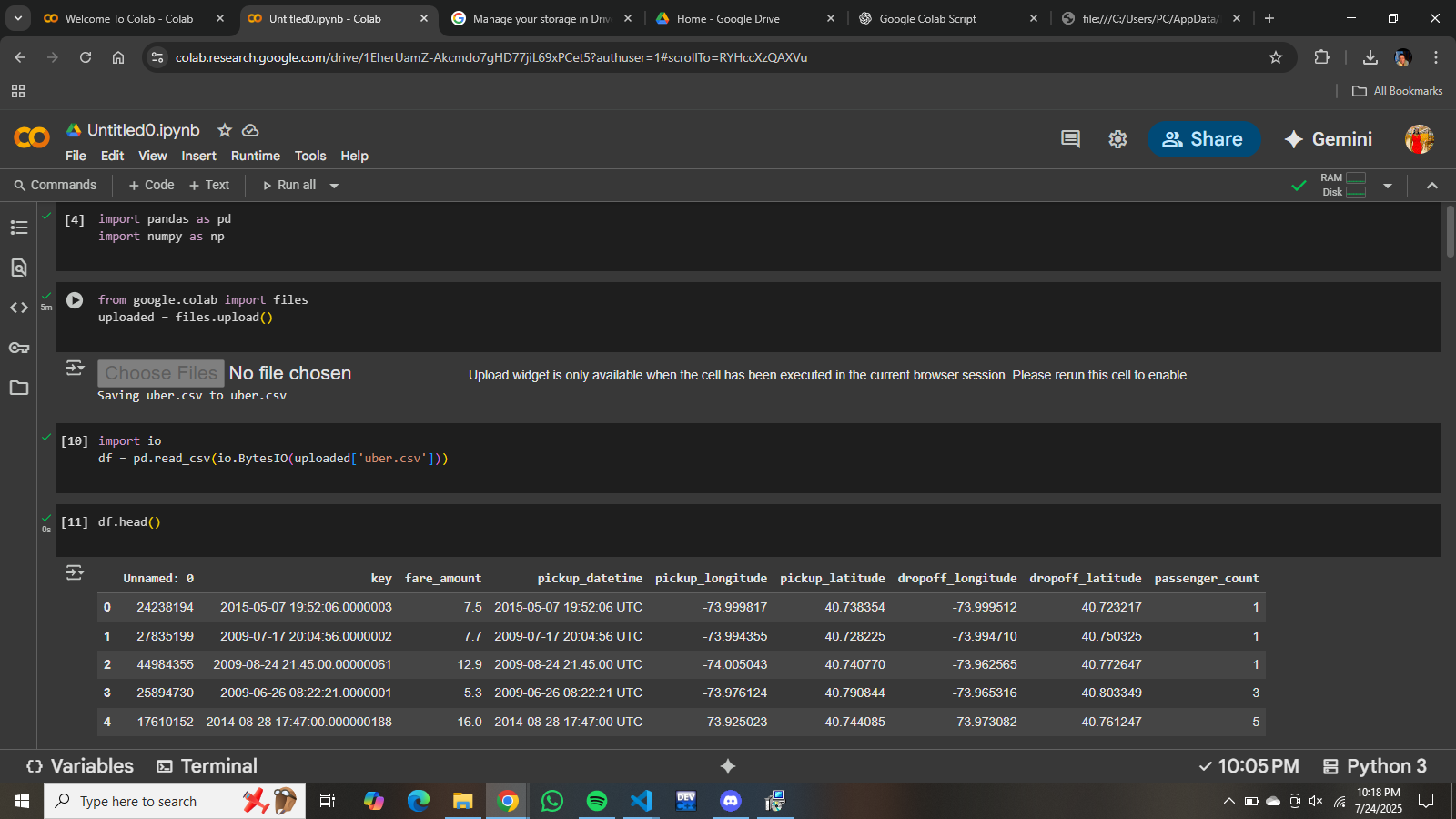
**Introduction**

This notebook demonstrates an exploratory data analysis (EDA) of Uber ride data using Python. The workflow covers mounting Google Drive, data loading, cleaning, feature engineering, and visualization to uncover patterns in ride frequency, fare amounts, and trip distances.

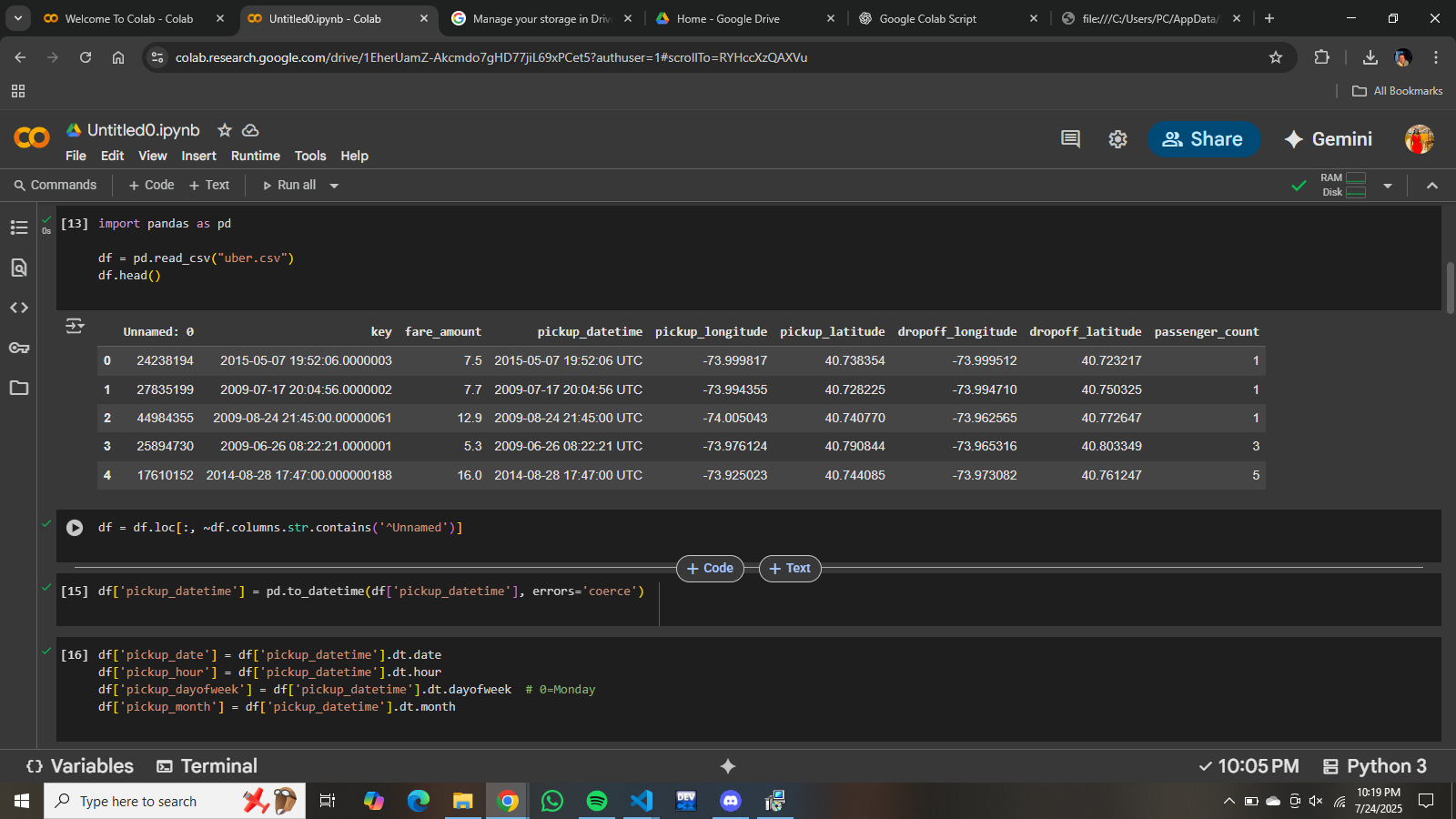
**Data Loading and Initial Exploration**



* **Upload and Read Data:**
* import pandas as pd
* import numpy as np

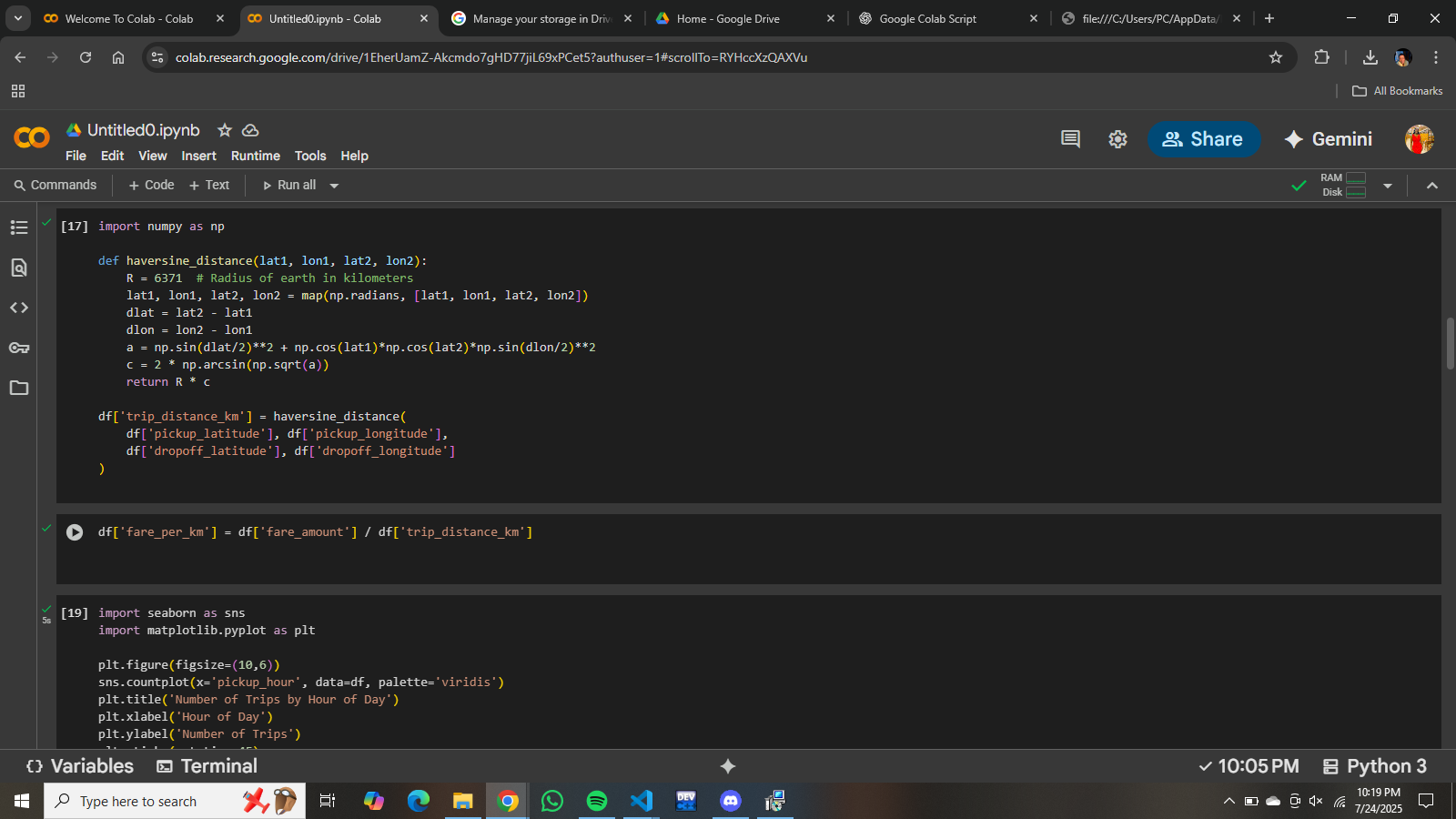
**Data Cleaning and Feature Engineering**

* Removed unnecessary columns (like "Unnamed: 0").
* Converted pickup\_datetime to datetime format.
* Extracted new features:
  + pickup\_date
  + pickup\_hour
  + pickup\_dayofweek
  + pickup\_month



**Calculating Trip Distance and Fare per Kilometer**

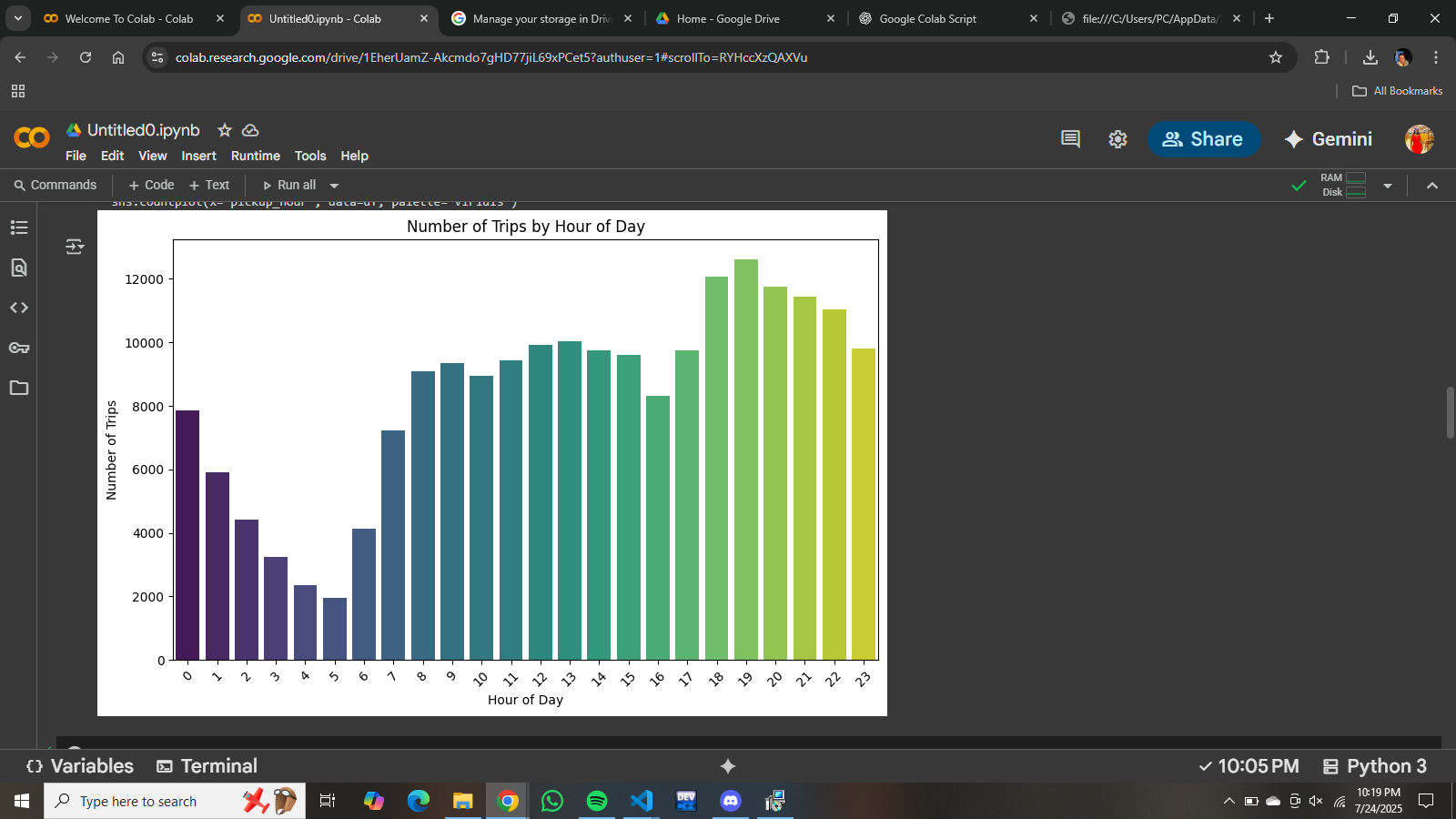
* Used the Haversine formula to calculate trip distance in kilometers.
* Added trip\_distance\_km and fare\_per\_km columns to the DataFrame.



**Visualizations**

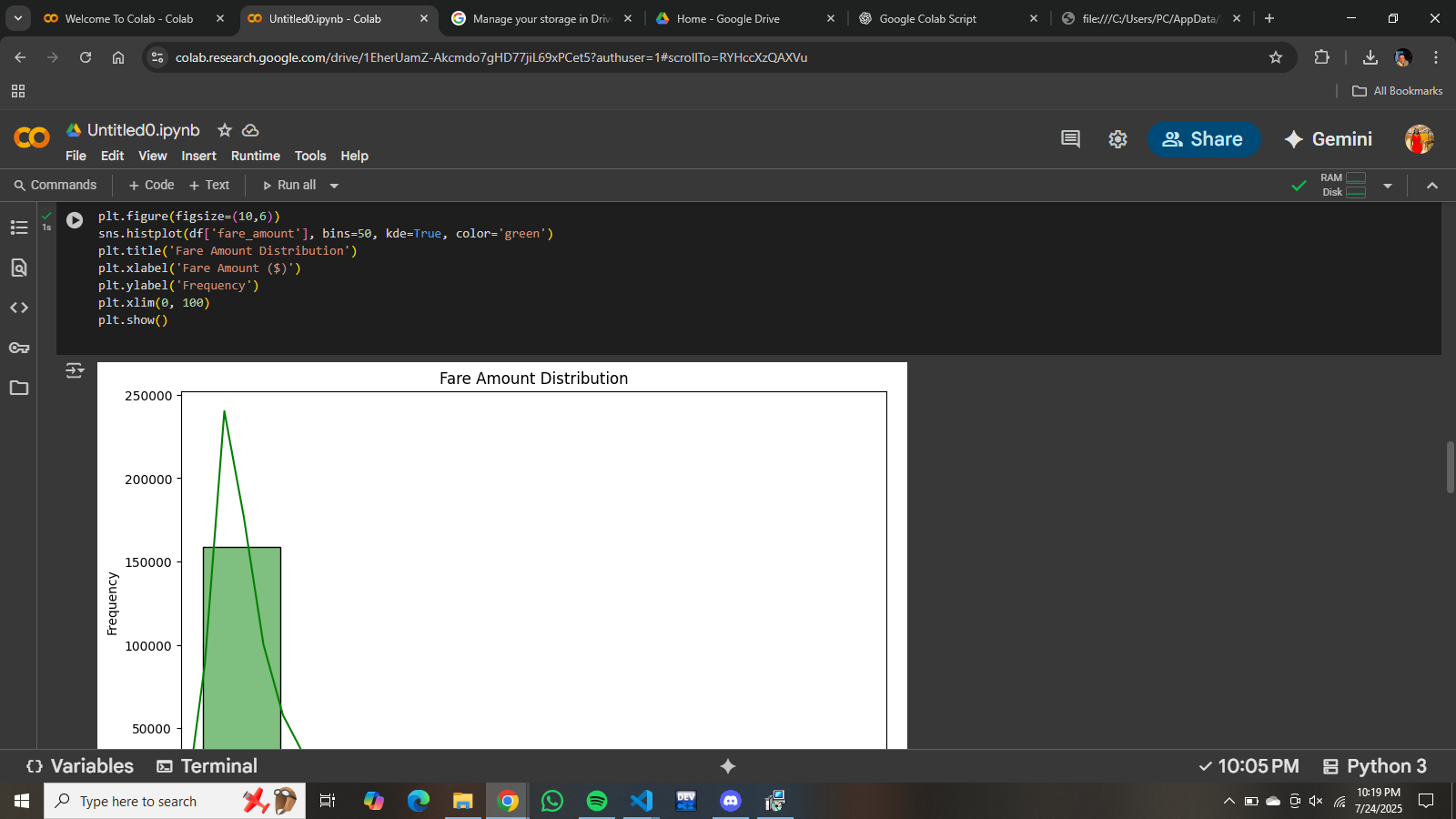
**1. Number of Trips by Hour of Day**

A bar plot shows the distribution of trips across each hour of the day. The highest number of trips occurs in the evening hours (18:00–21:00), with a noticeable dip during early morning hours.



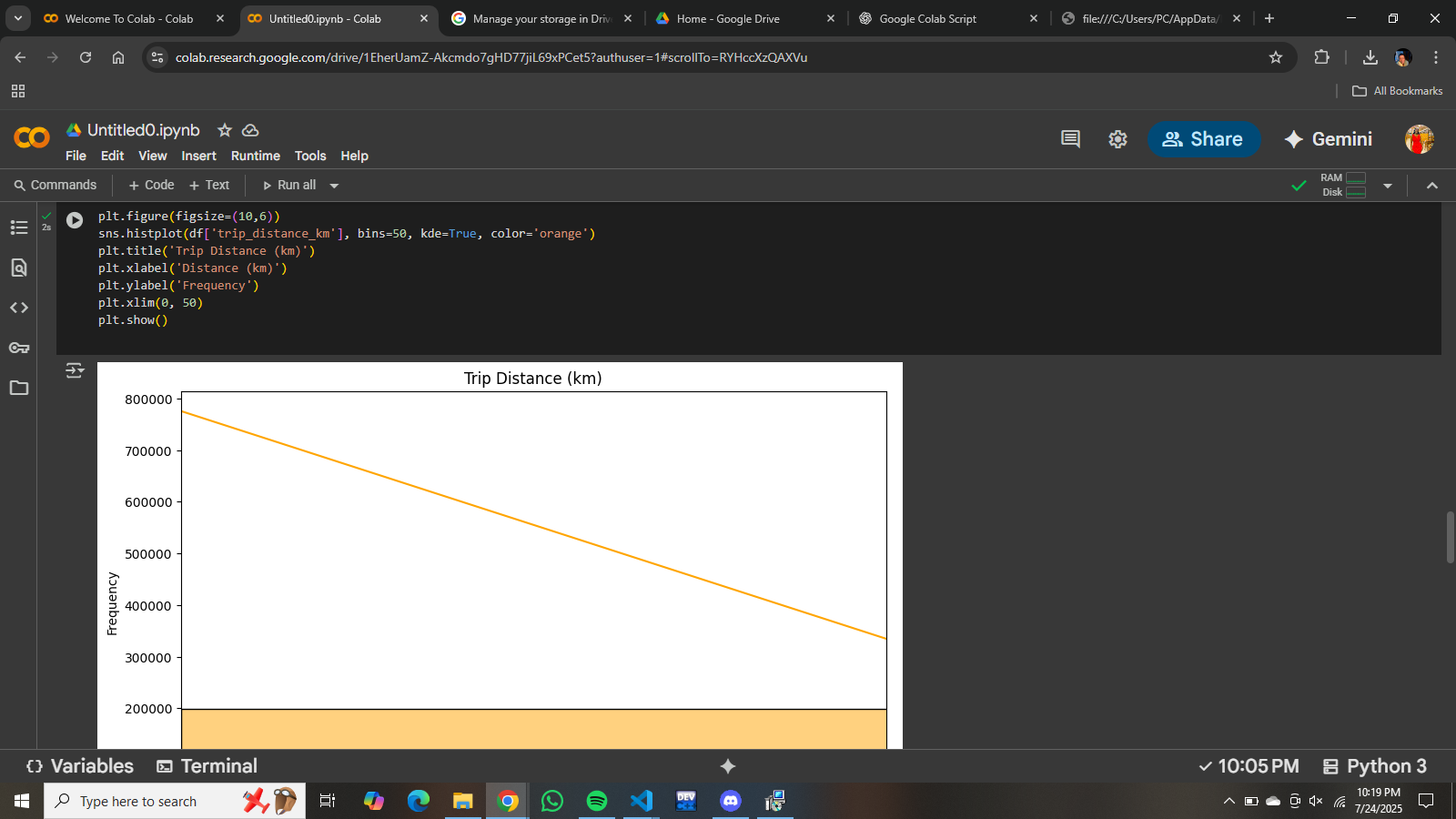
**2. Fare Amount Distribution**

A histogram (with KDE) displays the distribution of fare amounts. Most fares are concentrated below $20, with a sharp drop-off after that.



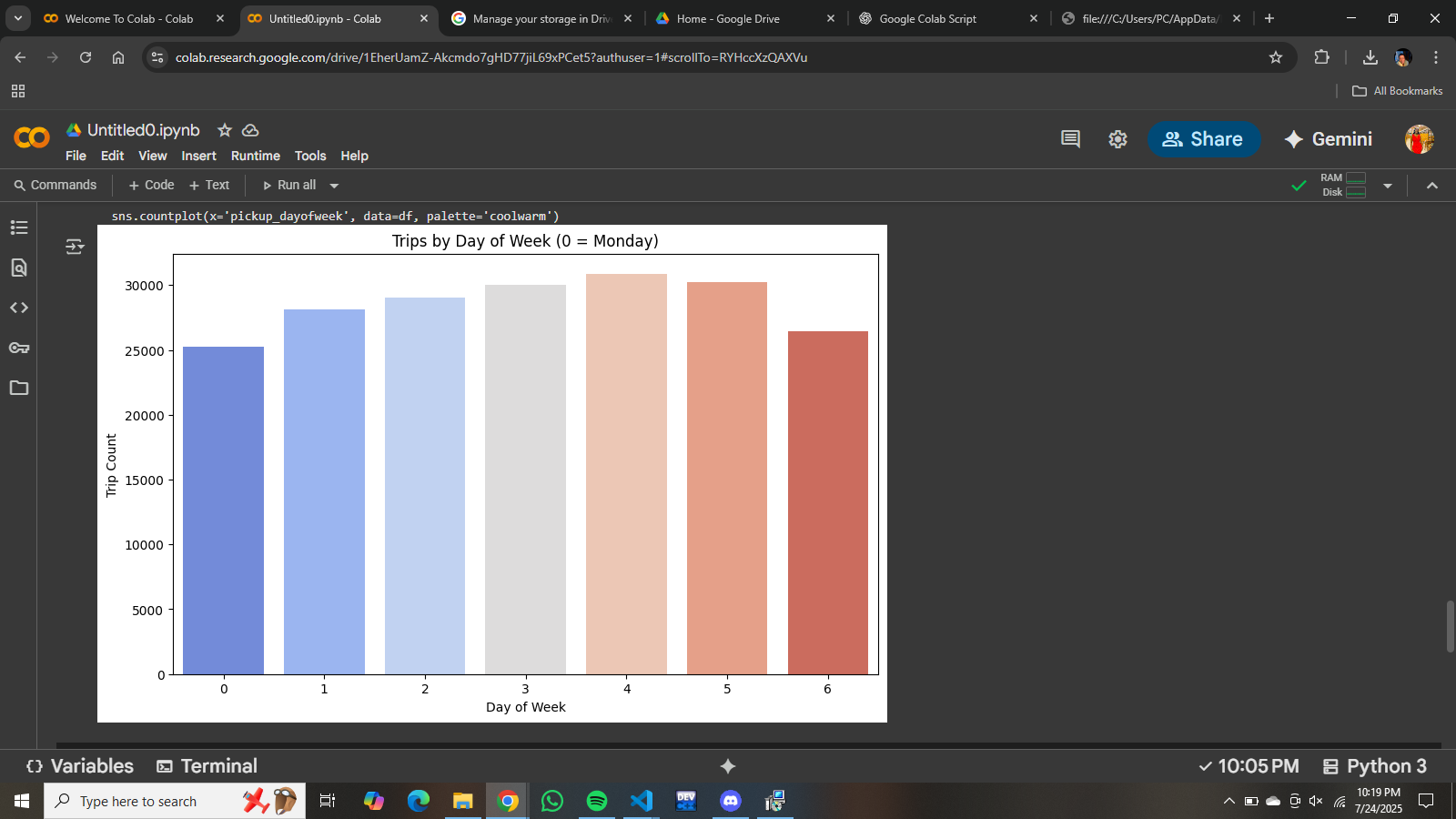
**3. Trip Distance Distribution**

A histogram (with KDE) shows that most trips are short, with the frequency decreasing as distance increases.



**4. Trips by Day of Week**

A bar plot shows the number of trips for each day of the week (0 = Monday). The number of trips is fairly consistent throughout the week, with a slight dip on Sundays.

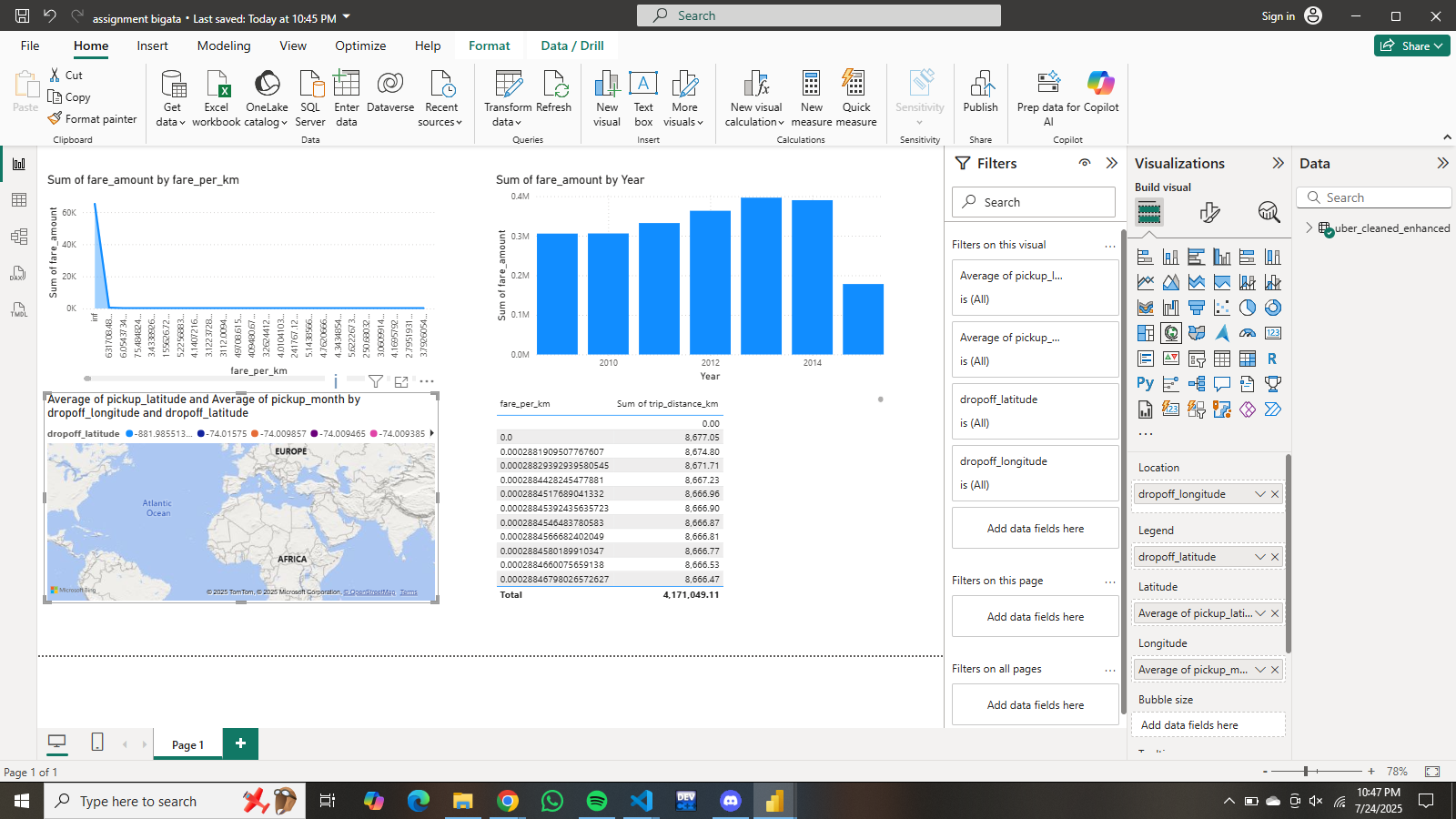


**5. Countplot Code for Day of Week**

The code for generating the day-of-week plot is also shown, confirming the use of Seaborn's countplot and the 'coolwarm' palette.

**Power BI Visualization**

The analysis also includes a Power BI dashboard, which provides interactive visualizations and insights based on the Uber ride data.



**Summary**

* The dataset was successfully loaded and cleaned.
* New features were engineered for deeper analysis.
* The trip distance and fare per kilometer were calculated.
* Visualizations revealed:
  + Most trips occur in the evening.
  + Most fares are under $20.
  + Most trips are short in distance.
  + Trip frequency is steady throughout the week, with a slight drop on Sundays.
* Power BI was used to create an interactive dashboard for further exploration.