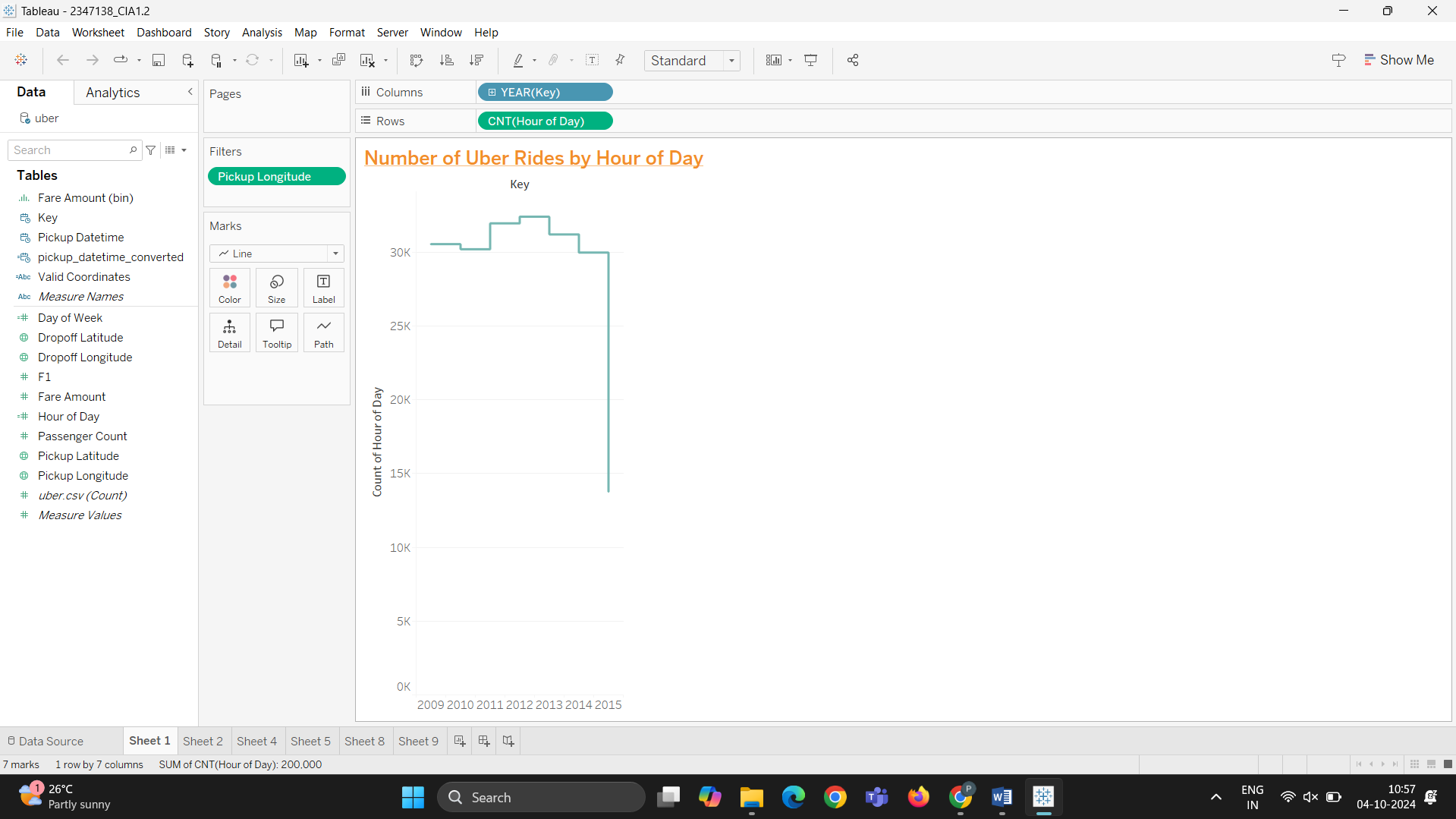
**DATA VISUALIZATION CIA**

**SUBMITTED BY-PRATHAM M (2347138)**

**Topic: Uber Transportation Analysis**

**Dataset link:** <https://www.kaggle.com/datasets/yasserh/uber-fares-dataset>

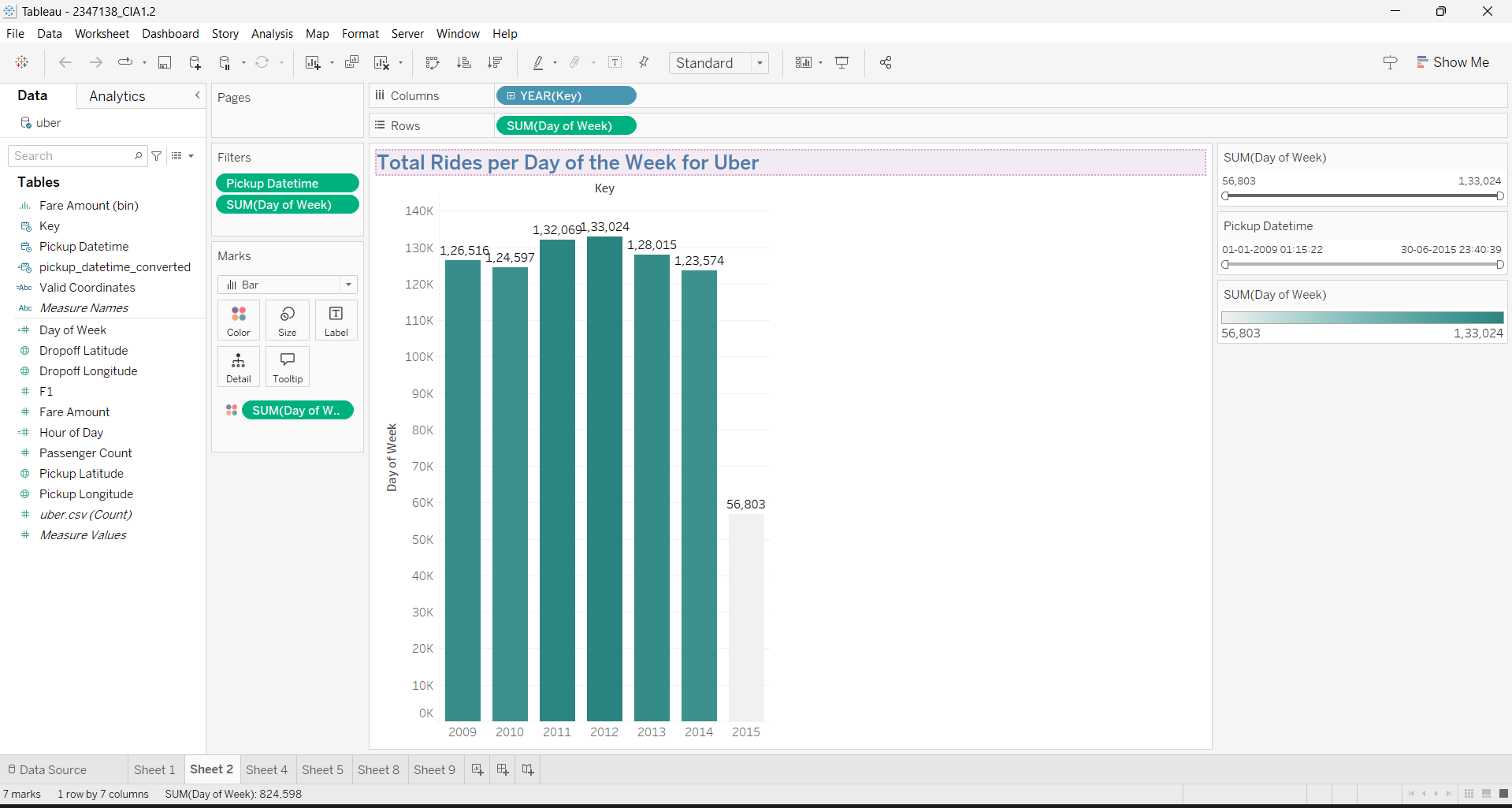
****

 **X-axis (Hour of Day)**: This shows different hours of the day when Uber rides were taken. It spans from midnight (0) to the end of the day (24).

 **Y-axis (Count of Key)**: This represents the number of Uber rides. Each point on the line represents the total number of rides taken during a specific hour of the day.

 **Key Observation**: The graph indicates how ride frequency fluctuates throughout the day:

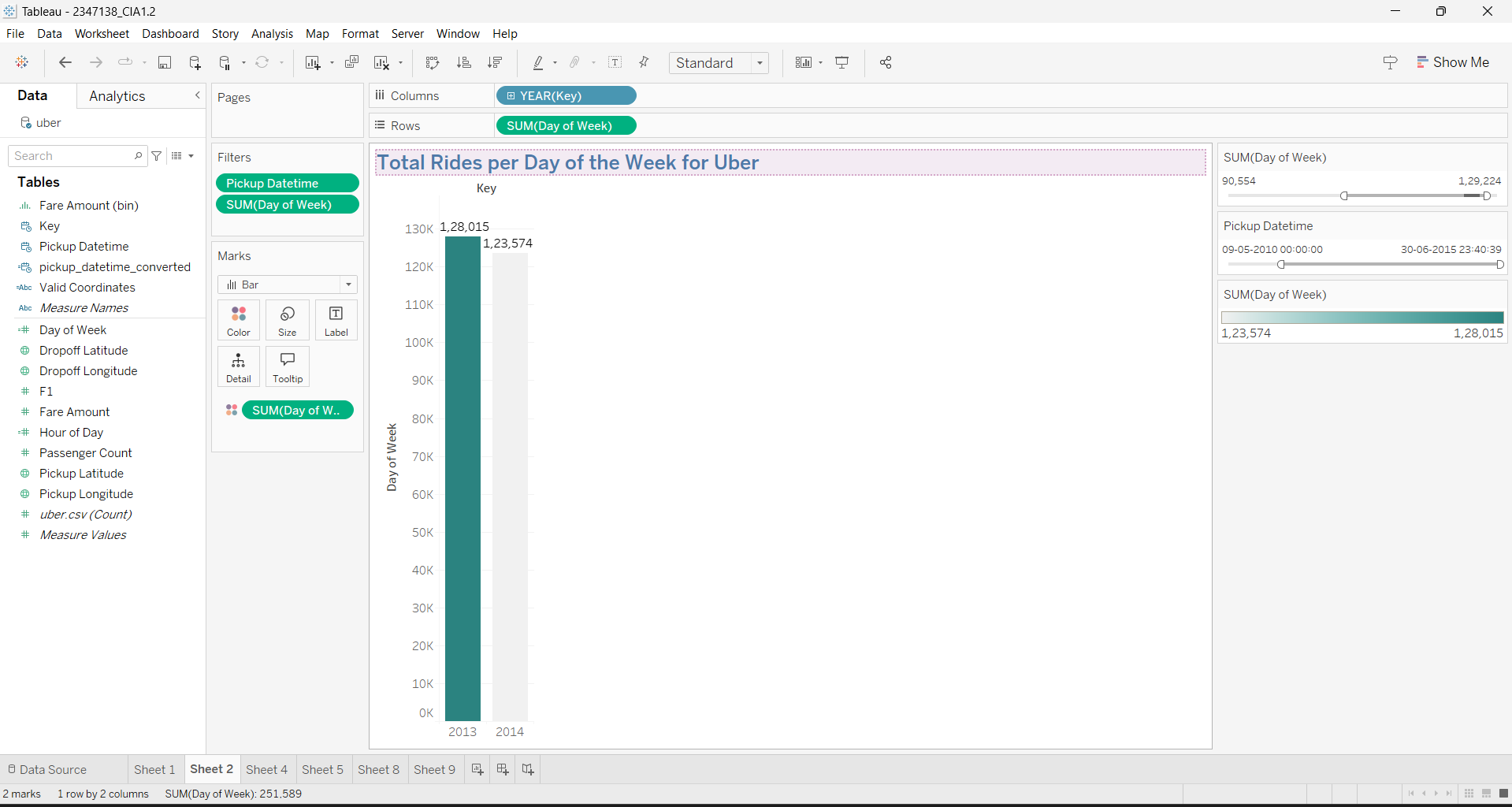
* **Rise and Fall**: There is a noticeable rise in the number of rides during the middle hours, suggesting more Uber rides are taken during certain times of the day, possibly correlating with high-demand periods (morning commute, afternoon rush).
* **Decline**: The number of rides drops sharply toward the end of the chart. This might indicate lower activity during late-night or very early morning hours.

****

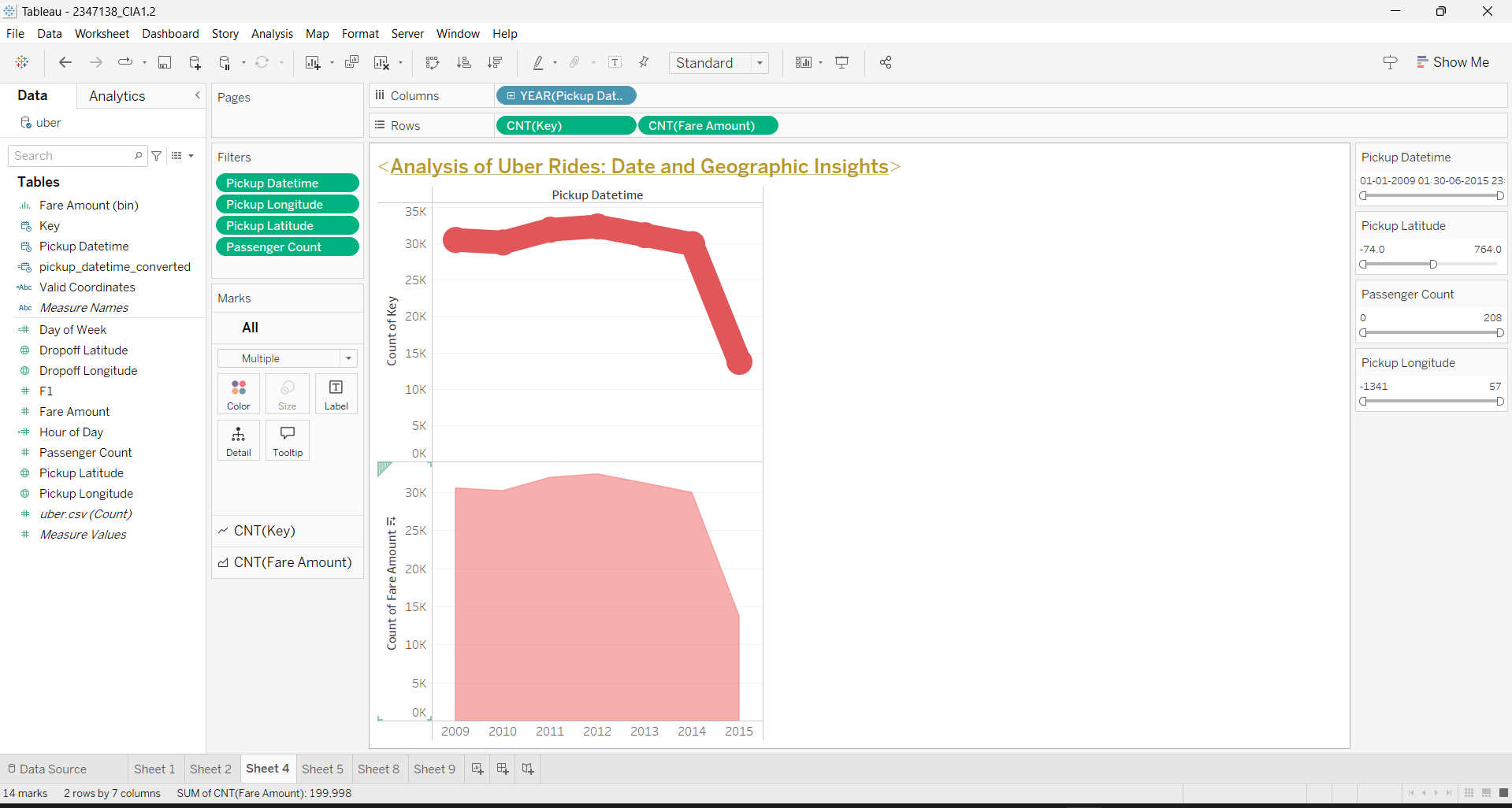
**Interpretation:**

1. **X-axis (Day of Week)**: This axis represents the days of the week, from Sunday to Saturday (0–6). It aggregates the number of rides for each day.
2. **Y-axis (Key - Total Number of Rides)**: The bars represent the total number of rides for each day of the week. The numbers above the bars (e.g., 1.2M, 1.3M) show the exact ride counts for each day.
3. **Total Ride Counts**:
   * **Sunday and Monday**: Both show the highest number of rides, with over 1.2 million rides each.
   * **Friday**: Slightly lower, with 1.18 million rides.
   * **Saturday**: The number of rides on Saturday seems to be the lowest (56,603), but this may be due to incomplete data for that day in 2015 or an issue with how the data is being visualized.
4. **Filter Context**:
   * The **Filter (Pickup Datetime)** on the right-hand side suggests that the rides are filtered for a specific range of pickup dates. It’s set between 30-06-2009 and 01-01-2015, so the data does not include rides after January 1, 2015.
   * The **SUM(Day of Week)** filter allows you to focus on total rides by day.

**After Applying the filter**

****

This chart titled **"Total Rides per Day of the Week for Uber"** shows the total number of rides in 2014, with approximately 1.23 million rides recorded. The visualization is filtered to display data only from 2014, excluding other years. The data suggests consistent Uber ride activity throughout 2014, but the absence of other years might limit broader trend analysis.

****

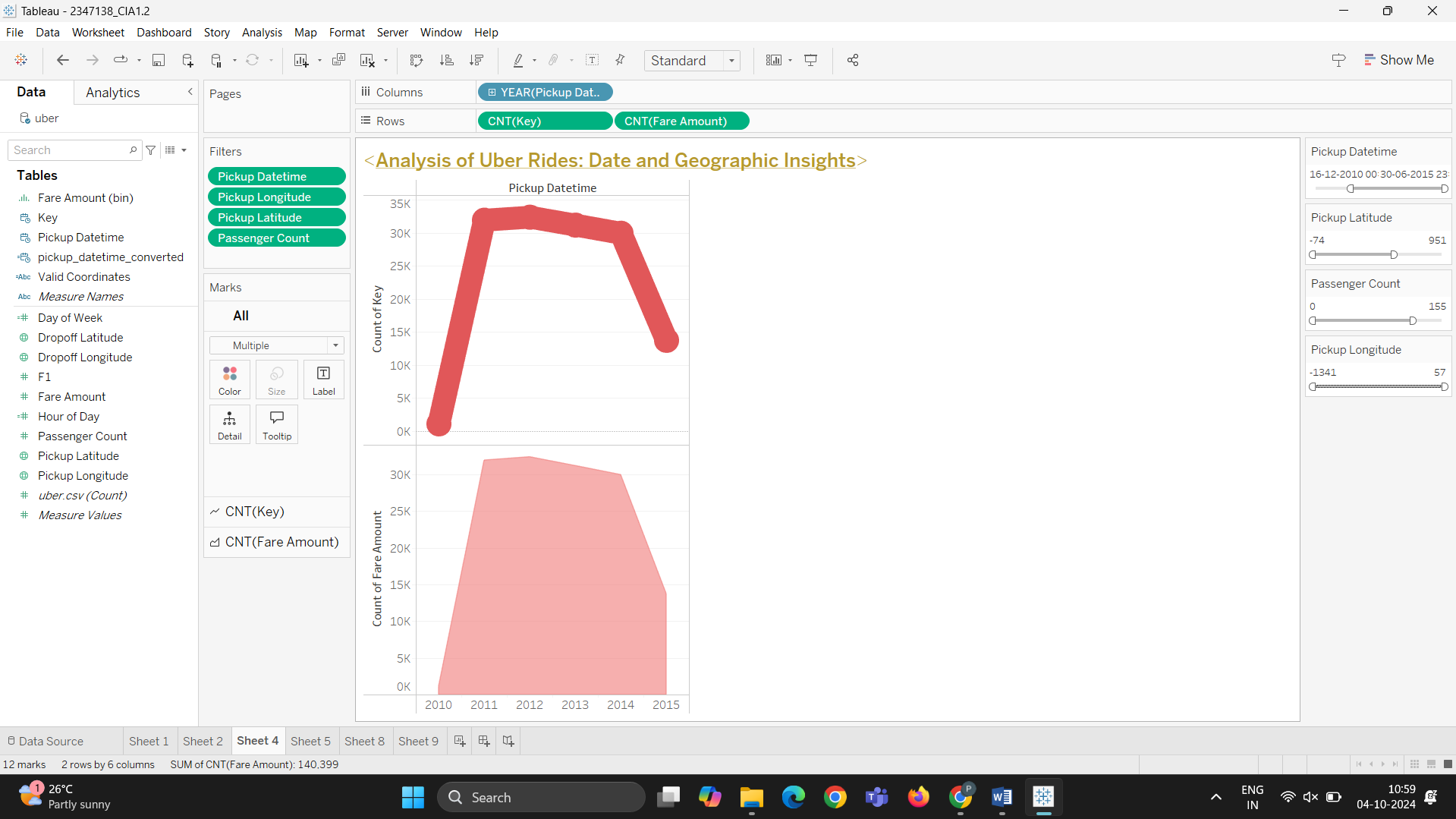
 **Time Range**: The x-axis represents the years (2009–2015), indicating that the data spans several years.

 **Metrics Displayed**:

* **Passenger Count (CNT(Passenger Count))**: Represented by the red solid line.
* **Fare Amount (CNT(Fare Amount))**: The pink shaded area represents the fare amount for Uber rides.

 **Trend**:

* **Passenger Count**: The passenger count starts around 2010 and increases gradually, peaking around 2014. However, there is a noticeable drop after 2014.
* **Fare Amount**: The fare amount generally follows a similar trend, increasing gradually, reaching a peak, and then declining sharply after 2014.

****

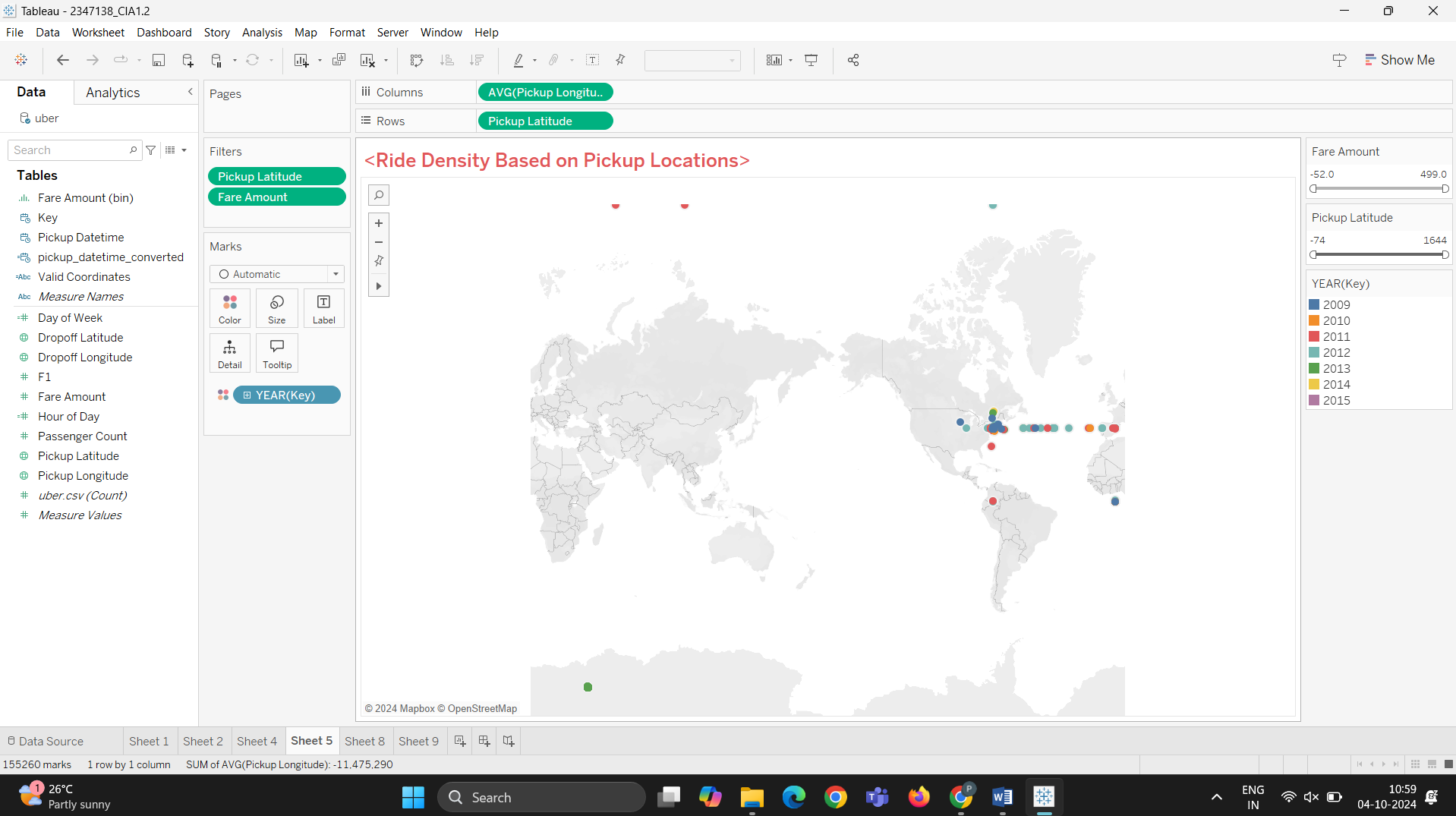
 **Passenger Count and Fare Amount**:

* The solid red line represents **Passenger Count** (CNT(Passenger Count)), while the shaded pink area represents the **Fare Amount** (CNT(Fare Amount)).
* Both metrics are charted across the years from 2011 to 2015 on the x-axis.

 **Time Range**: The x-axis now focuses on a shorter range (2011–2015), compared to the previous version.

 **Trend**:

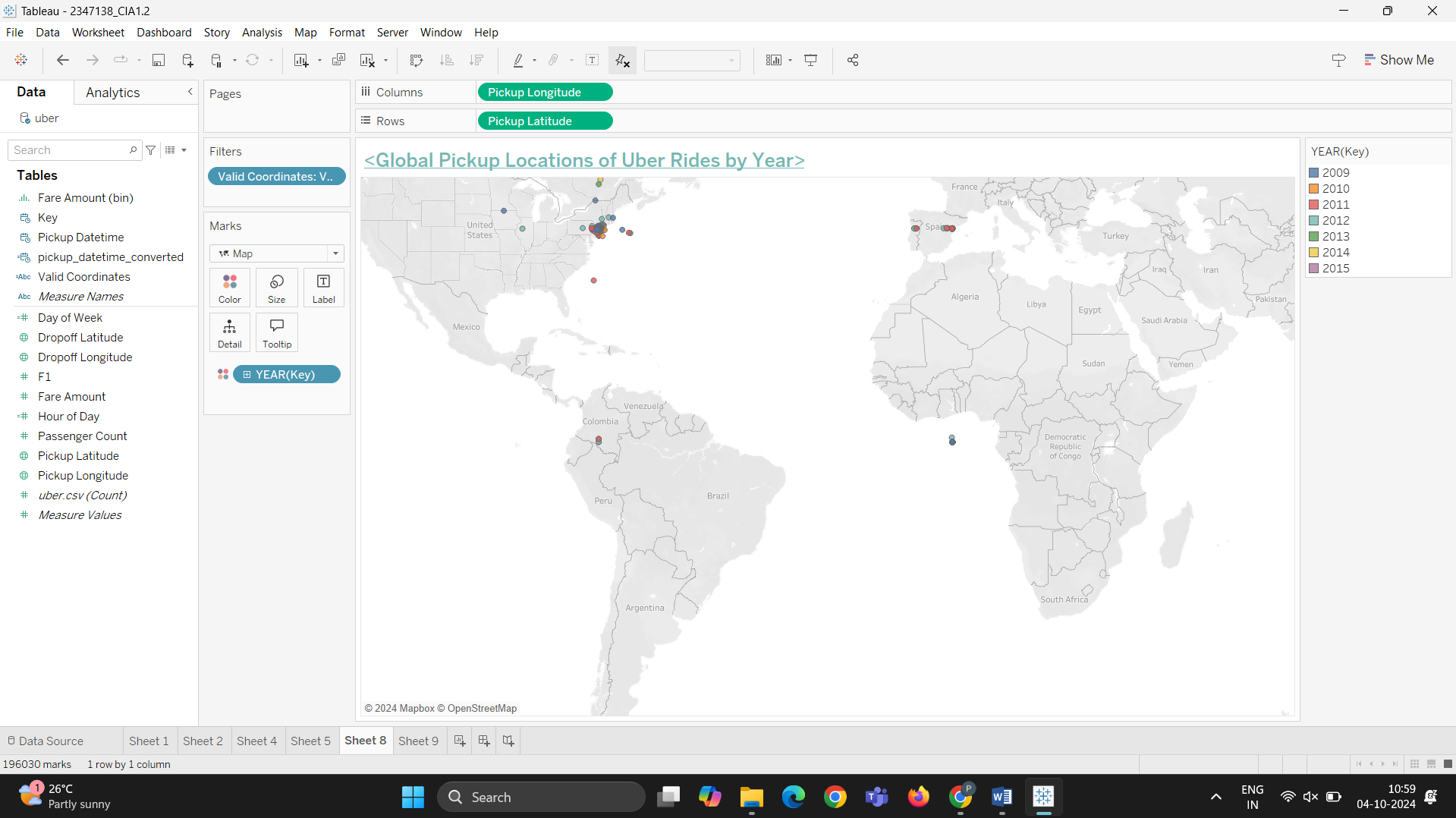
* **Passenger Count**: It rises steeply from 2011, peaks around 2013–2014, and then sharply declines after 2014.
* **Fare Amount**: Similarly, the fare amount follows a consistent pattern, rising in sync with the passenger count, peaking, and then declining.

****

1. **Map View**:
   * The visualization shows a global map where Uber rides are plotted according to their pickup locations.
   * The data points are concentrated in certain regions, particularly in **North America**, especially the **United States**. There are a few points in **South America** as well.
2. **Color-Coded Year Data**:
   * The data points are color-coded by year (2011–2015), which is displayed in the **legend** on the right.
   * Each dot represents a ride's pickup location, with different colors corresponding to the year of the ride.
3. **Fare Amount**:
   * There’s a **Fare Amount** metric, though it appears to be inactive or not clearly visible in this specific visualization.

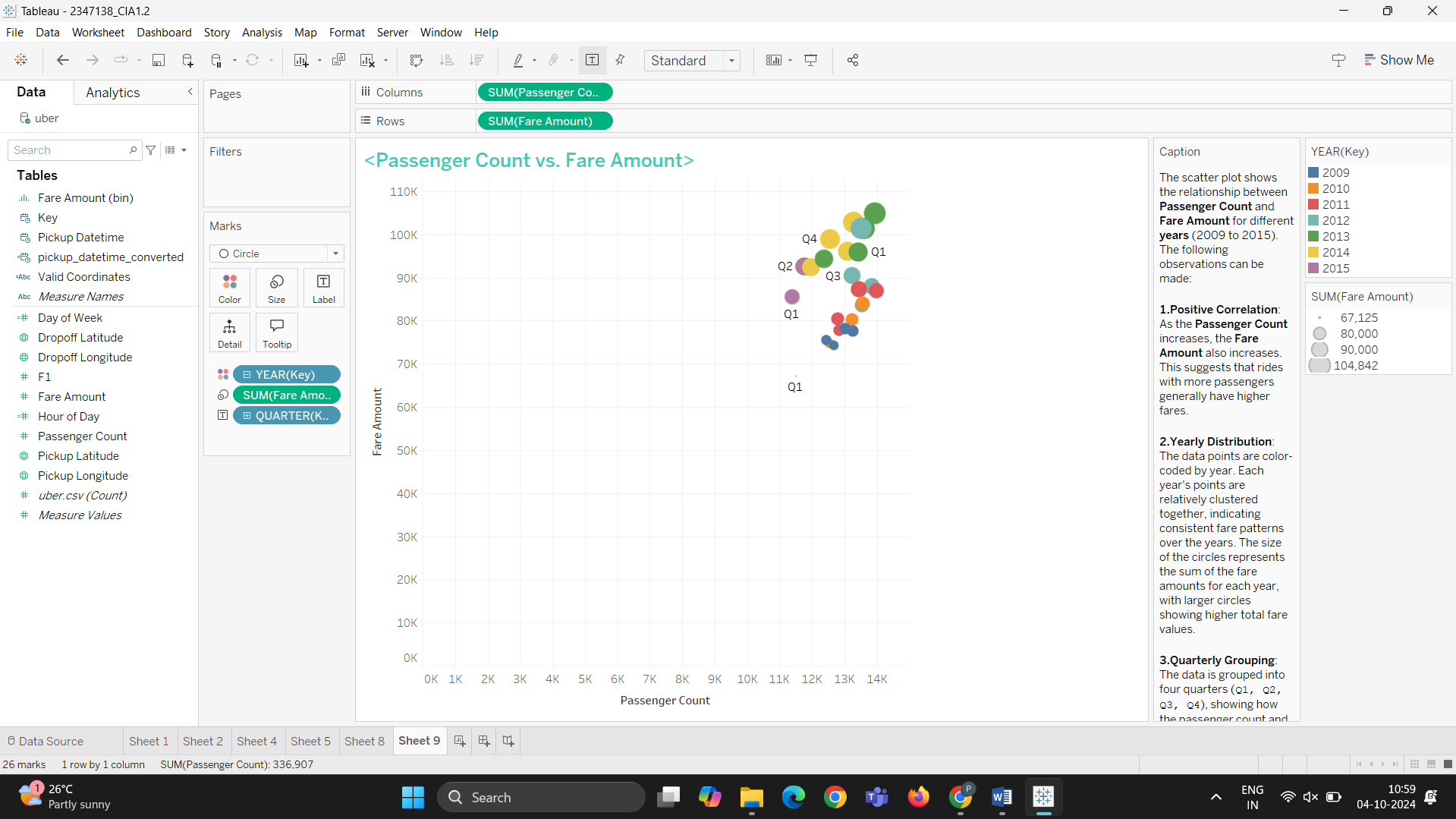
**Observations:**

* **High Ride Density in the US**: The majority of the Uber rides seem to be concentrated around the **eastern United States**, possibly in cities like New York or nearby urban areas.
* **Low Global Distribution**: Uber rides appear to have limited pickup locations outside the U.S., with sparse points in **South America** and very few in other parts of the world. This suggests that the dataset focuses primarily on U.S. ride data.

****

 Most pickup locations seem to be concentrated in the **United States**, particularly around the **East Coast** and parts of **South America**.

 There are also some locations in **Western Europe** (e.g., **France**) and **Africa**.

****

1. **Positive Correlation**:
   * As the **Passenger Count** increases, the **Fare Amount** also tends to increase, indicating a general trend where rides with more passengers typically have higher fare amounts.
2. **Yearly Distribution**:
   * The data points are color-coded by **YEAR(Key)**, representing different years (2009 to 2015).
   * The spread of the points for each year is relatively clustered, suggesting a consistent trend in the fare amount and passenger count across different years.
   * The size of each circle represents the **SUM(Fare Amount)** for the year, with larger circles indicating higher total fare values.
3. **Quarterly Grouping**:
   * The plot includes quarterly divisions (Q1, Q2, Q3, Q4), showing how the relationship between **Passenger Count** and **Fare Amount** varies across different quarters within the same year. For instance, Q1 to Q4 for each year show varying data points.

The chart includes a caption summarizing the key insights:

* **Positive Correlation** between passenger count and fare amount.
* **Yearly and Quarterly Trends**: Data is grouped by year and quarter to show specific patterns over time.