Name – Shreya Ganjewar

Roll no. – CS2-64

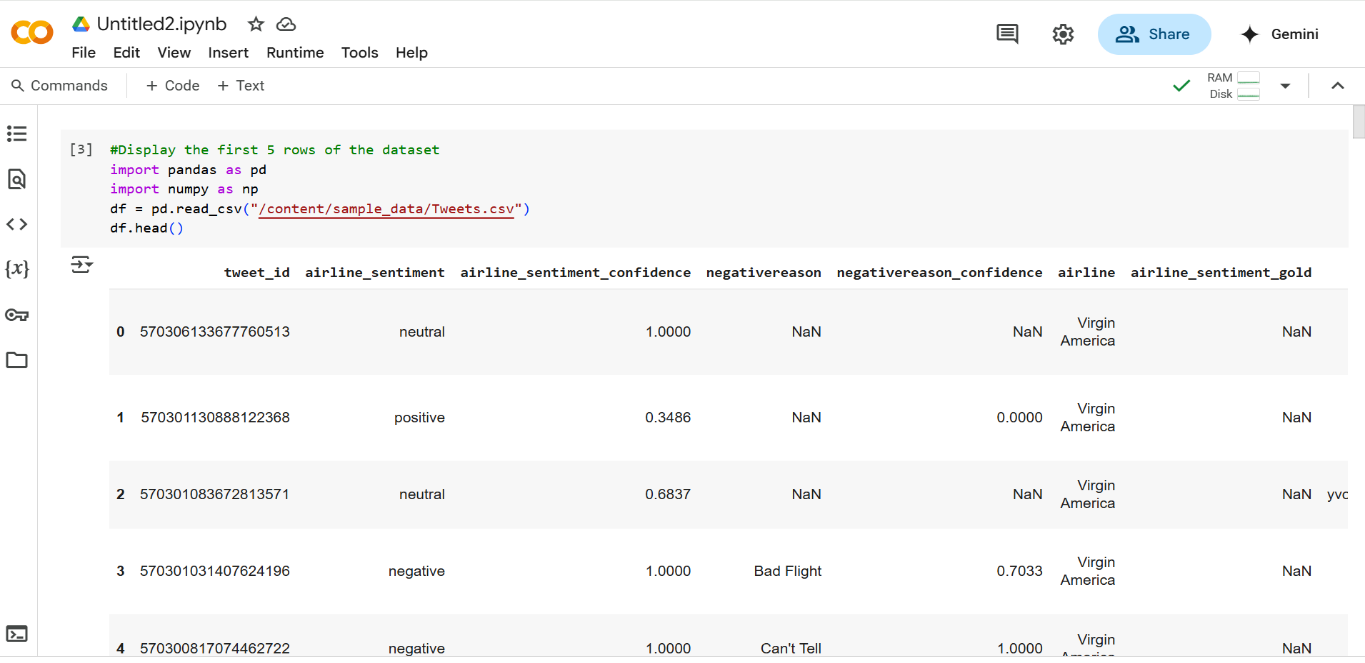
PRN – 202401040317

**EDS Theory Activity 1**

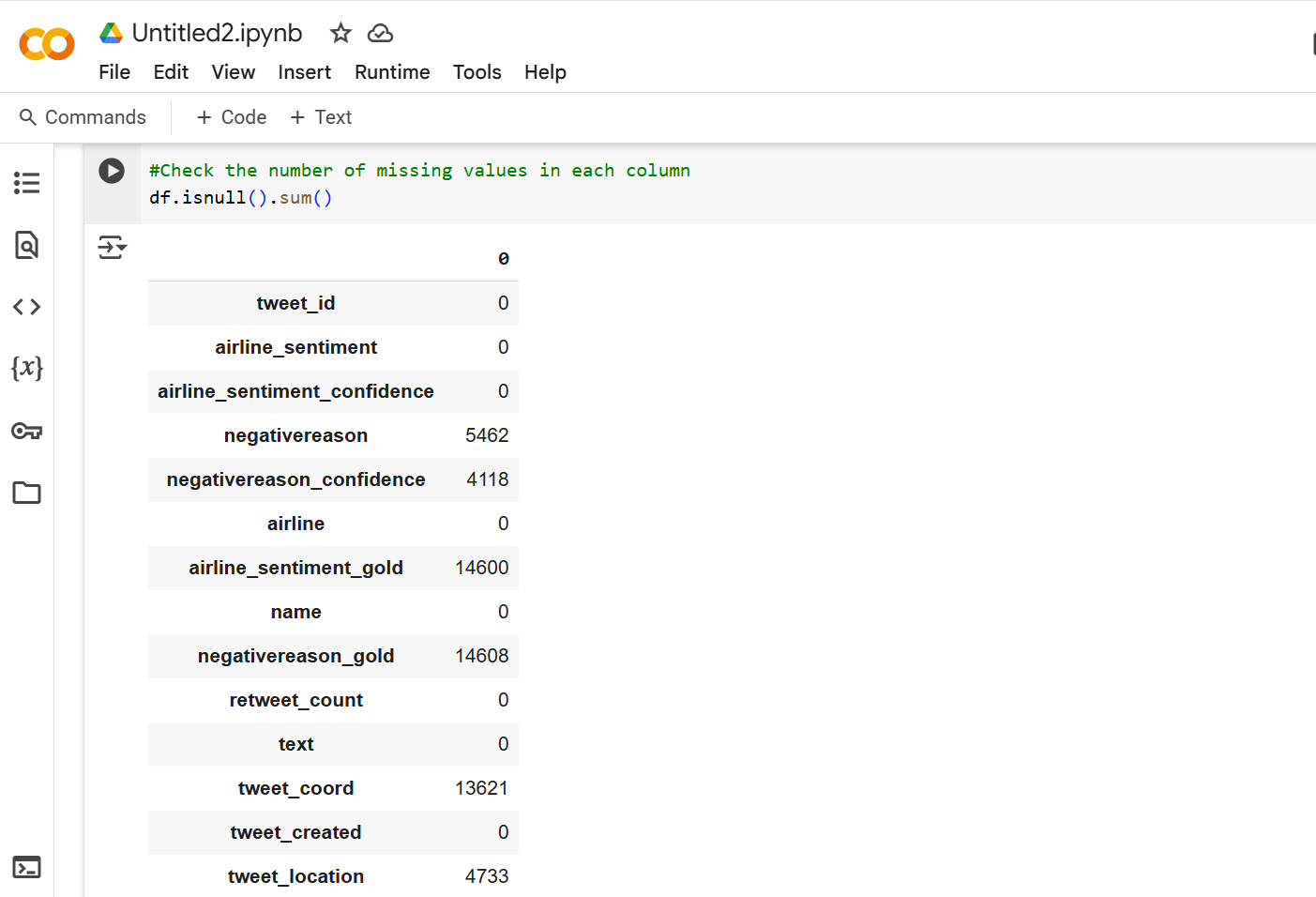
Dataset : Twitter US Airline Sentiment

**Problem Statements-**

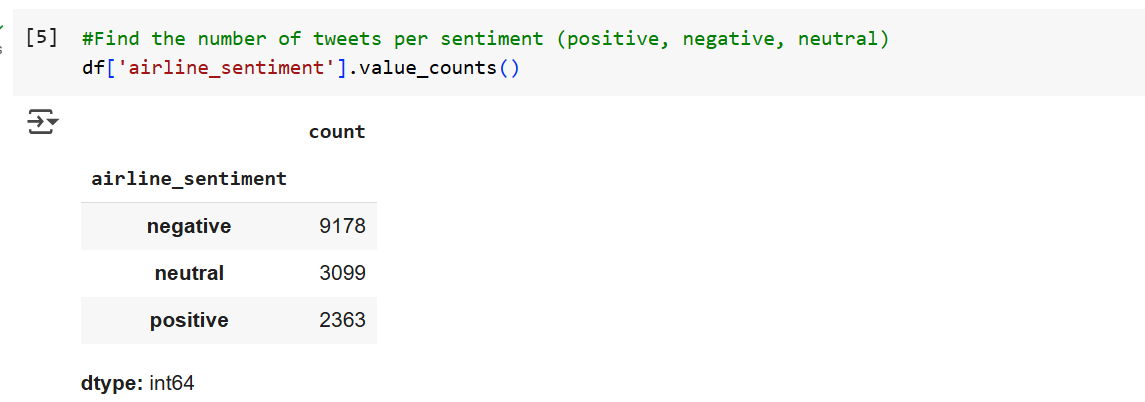
1. Display the first 5 rows of the dataset



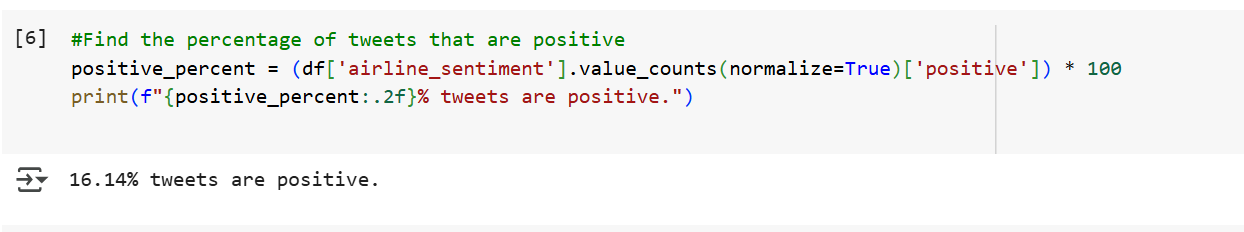
1. Check the number of missing values in each column



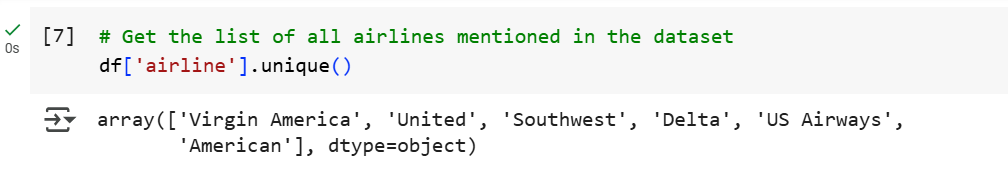
1. Find the number of tweets per sentiment



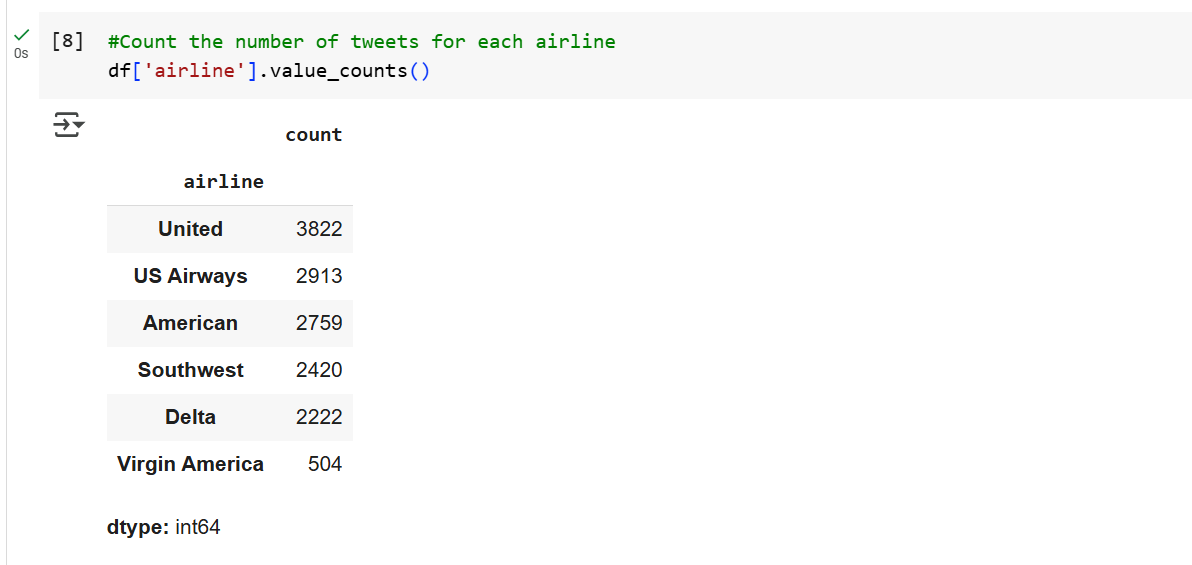
1. Find the percentage of tweets that are positive



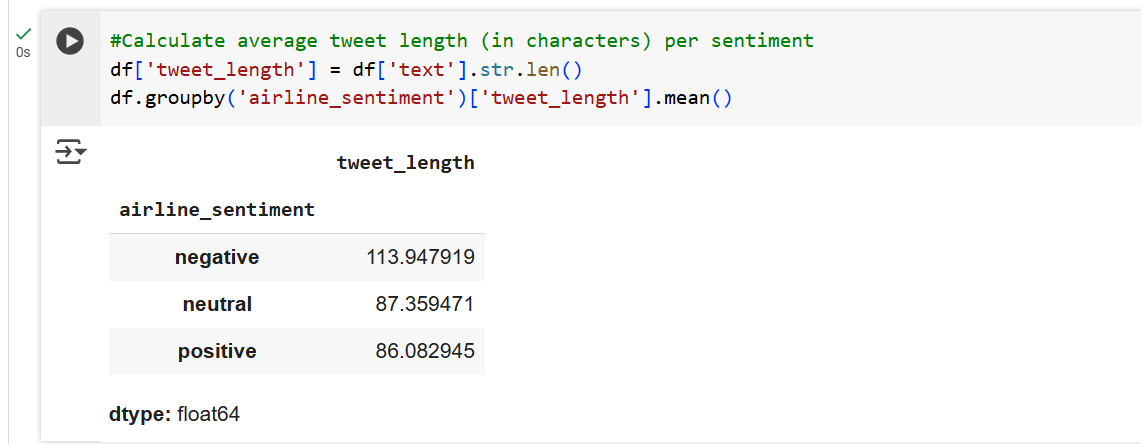
1. Get the list of all airlines mentioned in the dataset



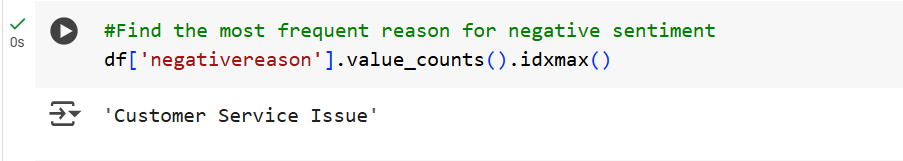
1. Count the number of tweets for each airline



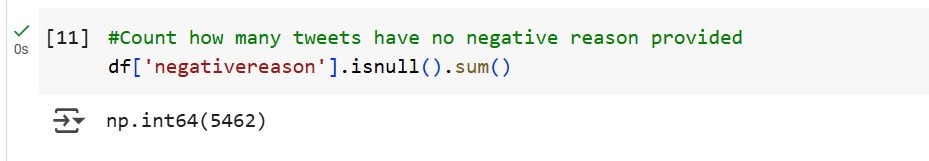
1. Calculate average tweet length per sentiment



1. Find the most frequent reason for negative sentiment

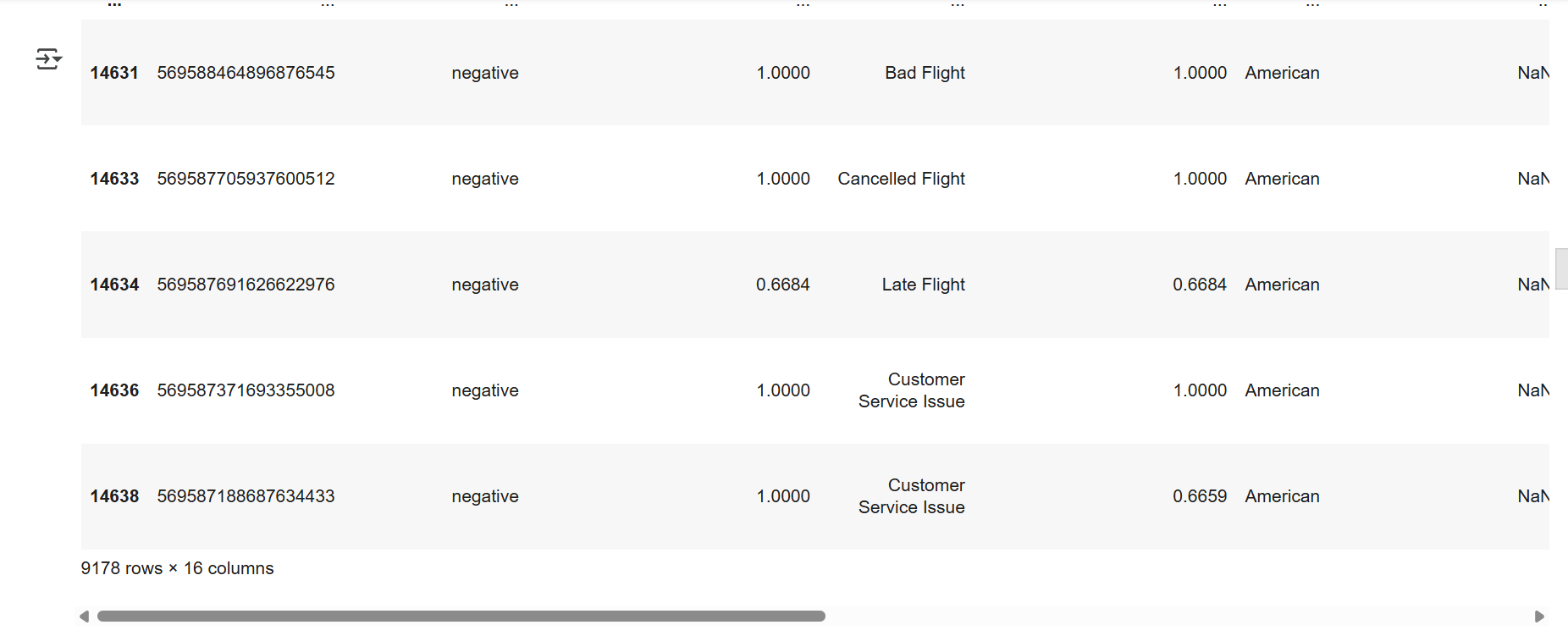


1. Count how many tweets have no negative reason provided



1. Filter only tweets that have a negative sentiment

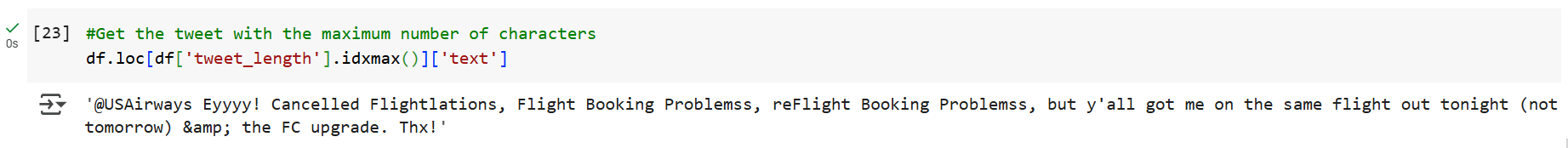




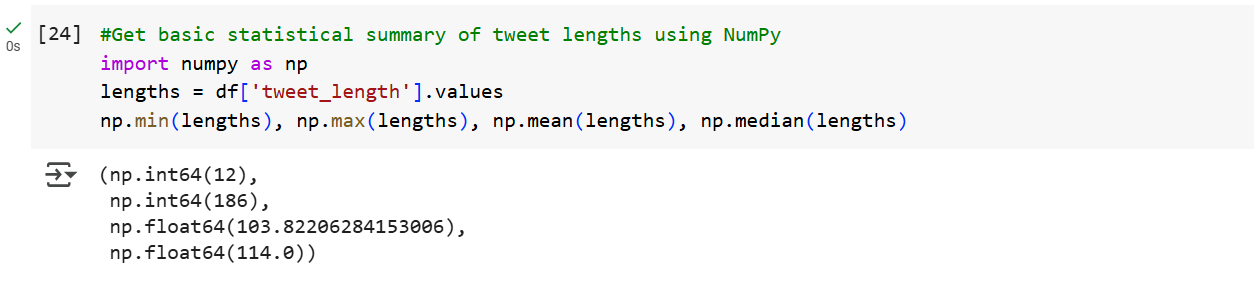
1. Group and count negative tweets per airline



1. Get the tweet with the maximum number of characters

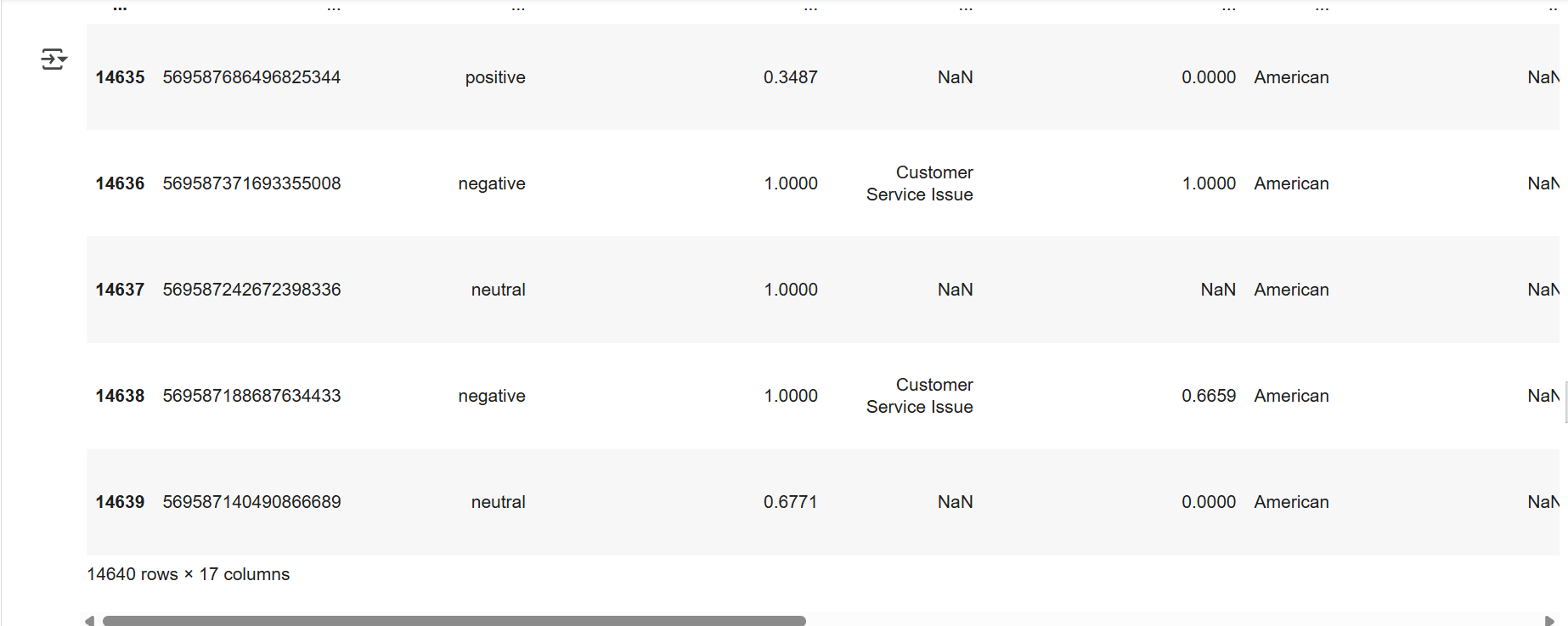


1. Get basic statistical summary of tweet lengths using NumPy

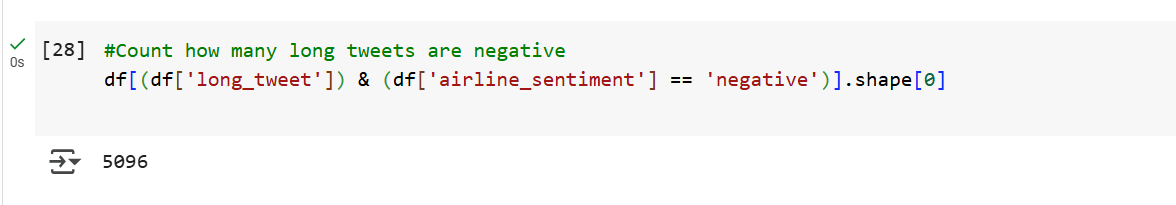


1. Create a column that flags tweets longer than 120 characters



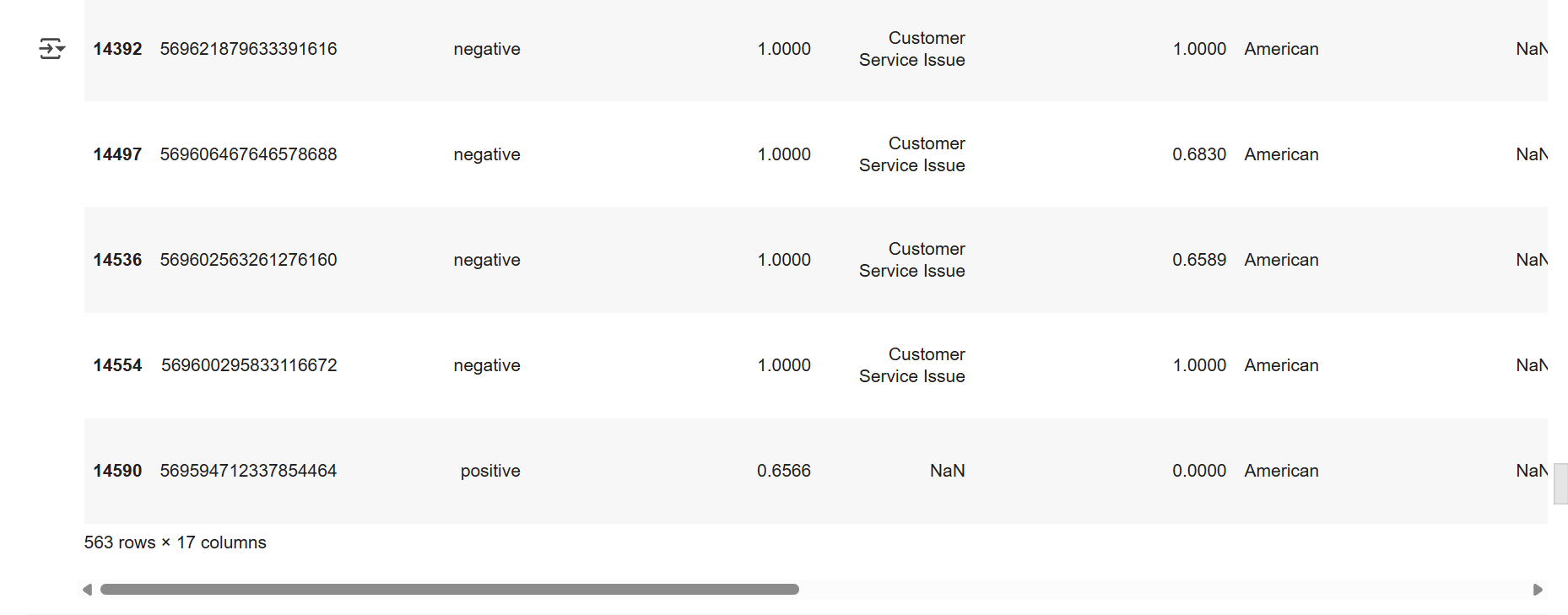


1. Count how many long tweets are negative

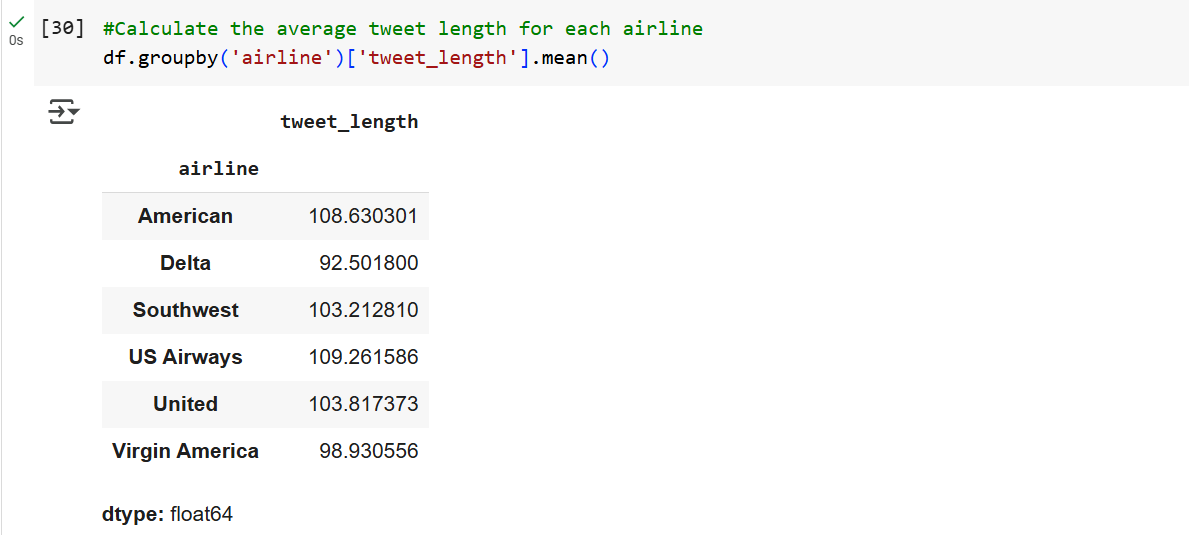


1. Extract all tweets that mention “customer service”

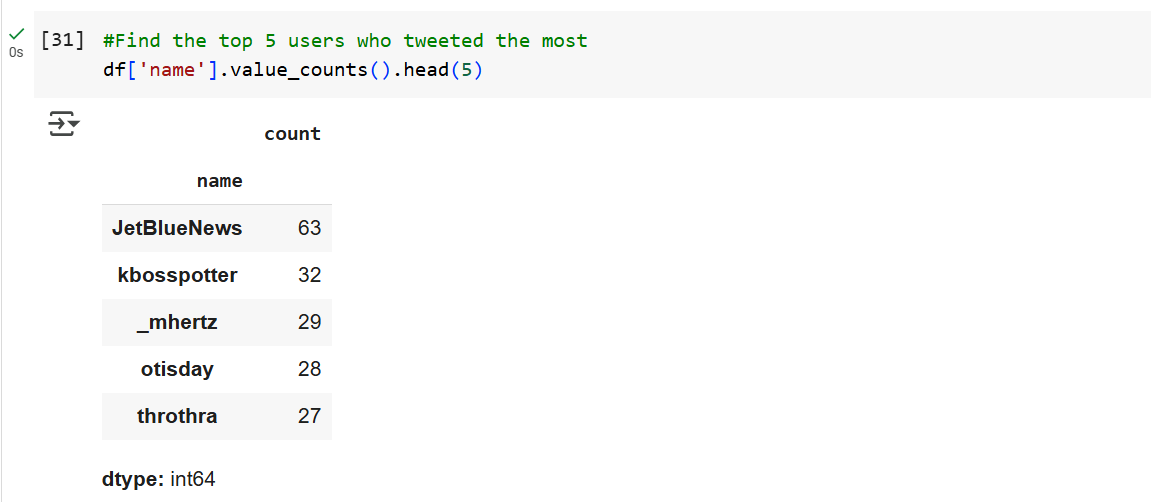




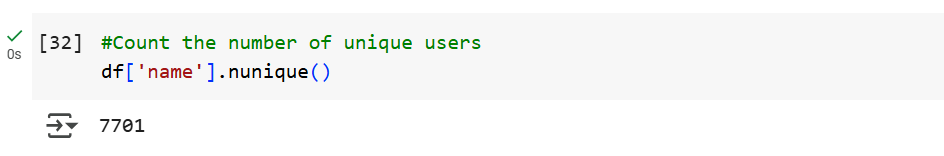
1. Calculate the average tweet length for each airline



1. Find the top 5 users who tweeted the most



1. Count the number of unique users



1. Plot the sentiment distribution for each airline



