- Which dataset you chose?

  I chose Flights dataset. The data is obtained from RITA

- Main findings from the exploratory data analysis, and how you chose the results to put in your explanatory analysis?

  After answering many questions, the following finding is listed:
    - Out of 5202096 flights from January 1988 to December 1988 around 50163 flights (0.96%) were canceled. In the following question, we will try to now the highest reasons that led to these cancellation
    - The ORD airport is the busiest airport with 5.2% of total flights While the DEN is less crowded airport. Overall, it's clear that the difference between the last six airports is small.
    - The number of flights per Day of Week graph where 1 (Monday) - 7 (Sunday), Friday (label 5) followed by Tuesday and Wednesday (label 2 and 3) has the maximum number of flights. Furthermore, January and December have the highest number of flights where February is the silent month of the year.
    - The relationship between both delays is positive
    - The January has the highest delayed flights for both departures and arrived opposite to April. Overall, on comparing to the start and end of the year the middle months is the less time of delay.
    - Out of the 14 Airlines the EA has the longest departure delayed. On the other hand, the PI and PA are considered as longest arrival delayed. Overall, it looks like some airlines have longer arrival delays than departure delays.