DS 4300: Large-Scale Storage and Retrieval
The Challenger Accident

Problem Description:

On January 28th, 1986, the space shuttle Challenger exploded approximately seventy seconds after lift-off, killing all seven astronauts on board. The accident was caused by the unusually cold temperatures reducing the resiliency of the rubber O-ring seals connecting each section of the solid rocket boosters. The loss of resiliency led to a gas leak at one of the O-ring joints and the eventual explosion.

How cold was it that day at Cape Canaveral?   In this assignment you will analyze temperature data for weather stations within 100 km of the NASA facility. Do your analysis in Scala, Spark and use Apache Zeppelin to create a spark notebook. When you complete your analysis, export your notebook to a .JSON file and submit the file for grading.

The notebook attached to this assignment is your starting point and will help you to organize your analysis. I've included scattered comments to help you. The basic approach is

1. Read the station and temperature data

2. Filter and clean up the data

3. Identify all weather stations within 100 km of Cape Canaveral (Note: not all of the stations necessarily recorded a temperature on any given day.)

4. Use inverse distance weighting to estimate the temperature at Cape Canaveral on January 28, 1986.

EXTRA CREDIT (10 Points)

How many degrees below normal was the temperature that day at Cape Canaveral? Use year-by-year historical data to compute the temperature and produce a plot of Year vs. Temperature for January 28th.

What to submit:

For this assignment, I'd like each student to do his or her own analysis. Discussions between students are ok but do your own work.

Submit your saved Zeppelin notebook file and please name the file: *lastname_hw5.json*