GLOBAL CLIMATE CHANGE: FACT OR MYTH ?

Data Analysis Weather Data Across the World

1914 TO 2014

AVERAGE TEMPERATURE DIFFERENCE ACROSS WORLD

30F   60F

WINTER SEASON

FALL SEASON

USA
SEASONAL TRENDS
TEMPERATURE AND PRECIPITATION
1900 - 2016

SUMMER SEASON

TEMP TREND LINE

TEMPERATURE

PRECIPITATION

SPRING SEASON
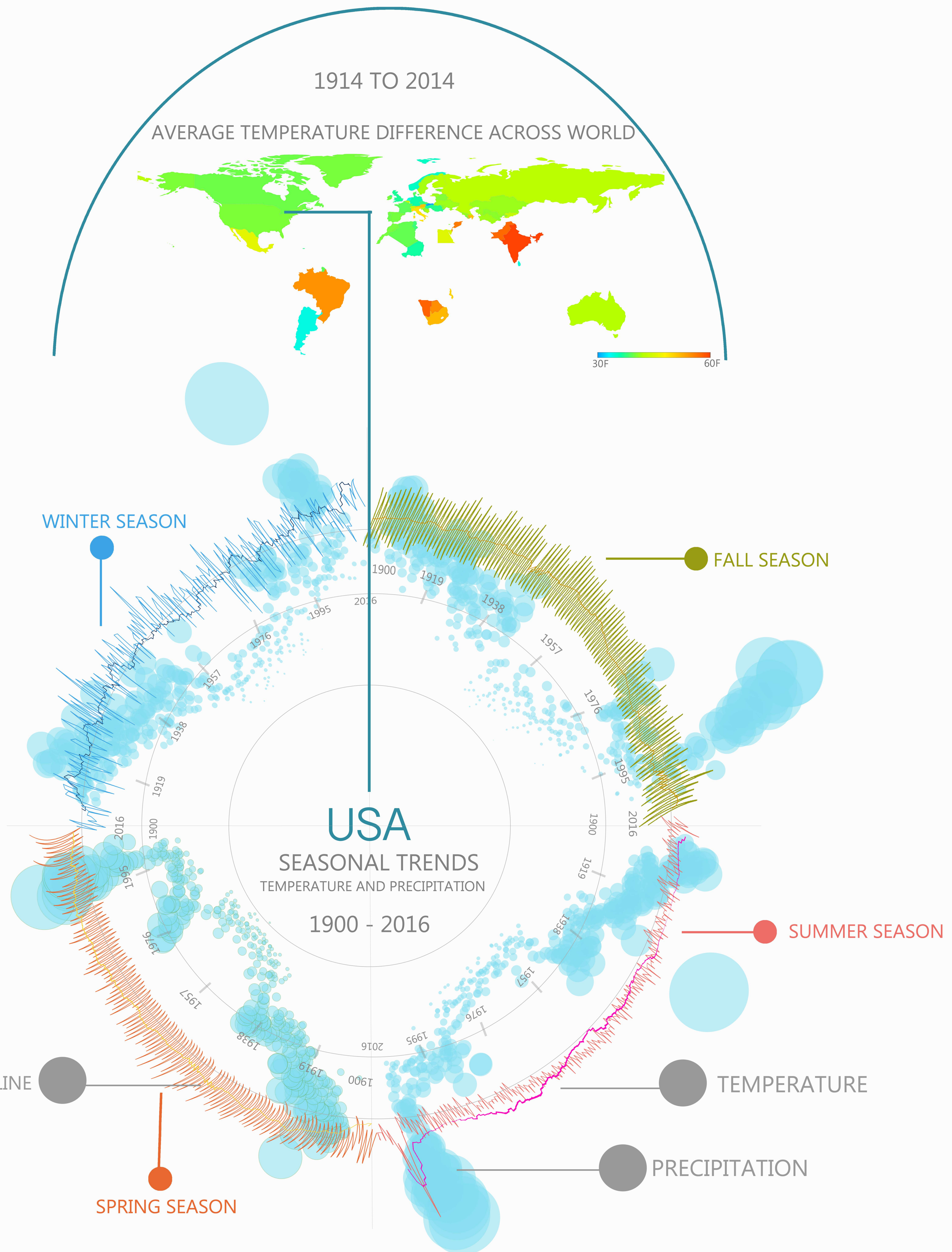
CHALLENGE:
DATA SIZE: 25GB
1000,790 Files from
Multiple Weather Stations

Download data
into Hadoop cluster

Use MapReduce
Calculate monthly
and Yearly Avgs
from daily data

Used Spark,
to create RDDs
and save
as a parquet file

Spark, SQL and Hive
Context to query
table returning .
csv files for analysis

Create Time Series and
Used Rolling Statistics
to identify Trend

Use Python to analyze data
and create visualizations

Created By - Avinash Kuncham, Alain Shema, Michael Spaulding, Mohit Singh
Data Source: ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily/

CLASS - IST 718 PROFESSOR - DANIEL ACUNA