

Comprehensive Analysis of U.S. 30-year Climate Normals [1991-2020]

Anbazhagan Naresh

June 19, 2024

Introduction

This document provides a comprehensive analysis of the 30-year climate normals data for key weather stations in the United States. The dataset, provided by the National Centers for Environmental Information (NCEI), includes monthly climate normals for temperature and precipitation for the period from 1991 to 2020.

Data Acquisition

The climate normals data were sourced from the NCEI website. The inventory and readme files were downloaded to identify key stations and understand the structure of the dataset. The complete dataset can be accessed [here](#). The inventory file, which contains a list of all stations and their metadata, additional documentation explaining the methodology and data structure is available [here](#).

Key Stations

There are over 15000 stations across the United States and for our convenience of processing this data we identified key stations based on their Global Surface Network (GSN) and Historical Climatology Network (HCN) status. The list of key stations is extracted from the provided inventory file available on the NOAA website. We identified a total of 60 key stations from the inventory file.

Data Cleaning and Aggregation

The inventory data were read and cleaned to extract the relevant information for the key stations. The downloaded CSV files for each station were then read, and the data were aggregated to compute monthly averages for temperature and precipitation. Missing values were handled by removing any incomplete records to ensure data quality.

Data Characteristics

This dataframe has 720 rows and 414 columns. The names of the columns and a brief description of each are in the table below:

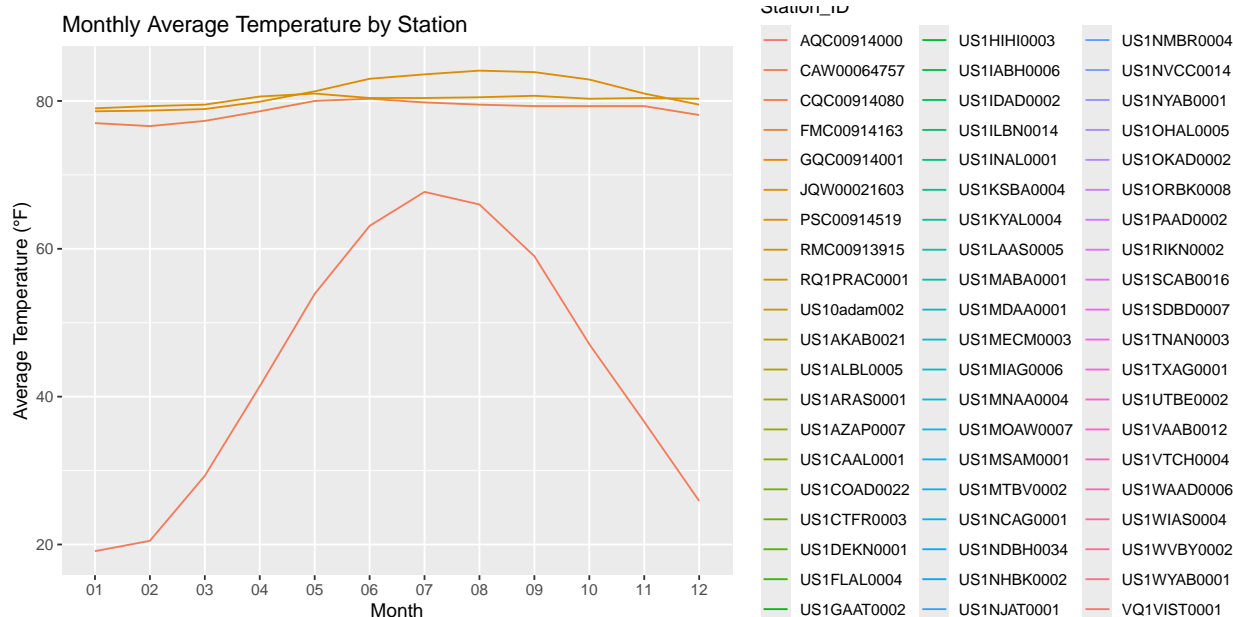
Summary Statistics

The summary statistics of the cleaned data provide an overview of the distribution of temperature and precipitation values across the key stations. The summary includes measures such as the minimum, maximum, mean, and median values, as well as the first and third quartiles for each variable. These statistics help us understand the central tendency and variability in the climate data.

Visualizations

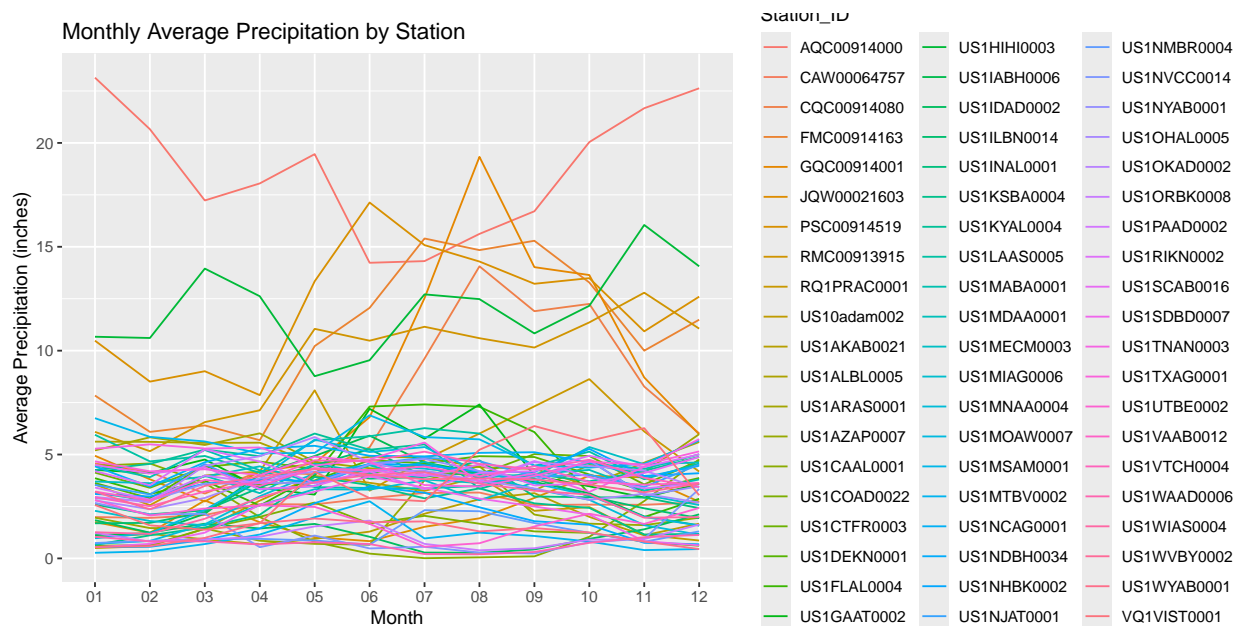
Monthly Average Temperature by Station

The line plot of monthly average temperature shows the variation in average temperatures across different stations over the year. This plot helps identify seasonal trends and compare the temperature profiles of various stations.



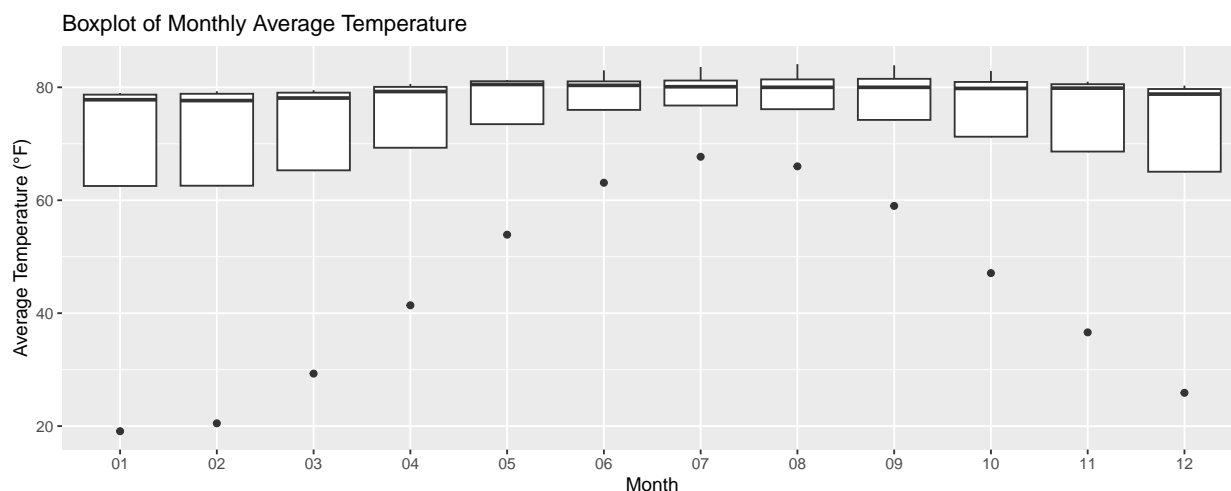
Monthly Average Precipitation by Station

The line plot of monthly average precipitation illustrates the variation in precipitation levels across different stations over the year. This plot helps identify periods of high and low precipitation and compare precipitation trends among stations.



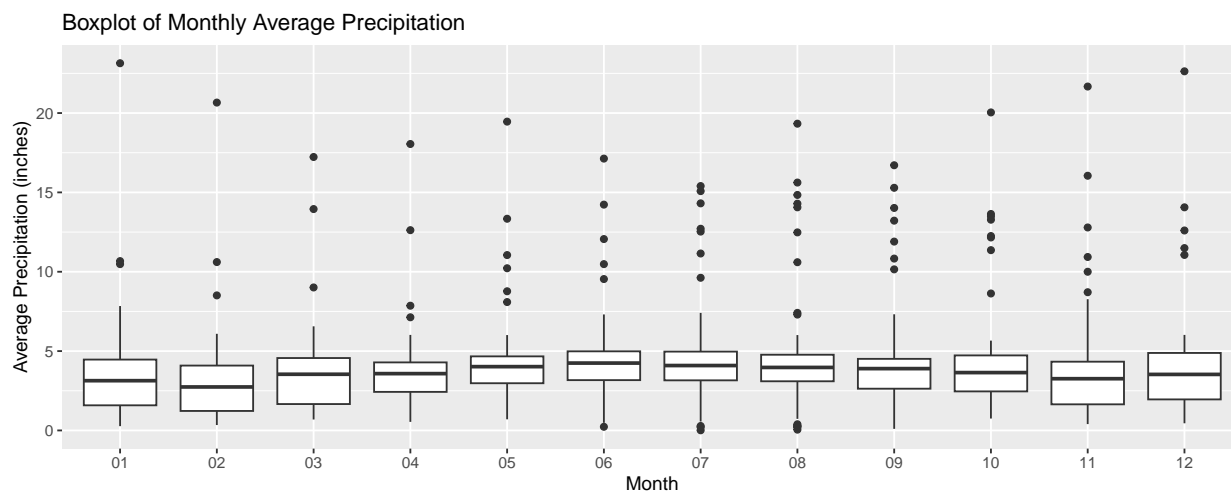
Boxplot of Monthly Average Temperature

The boxplot of monthly average temperatures reveals several outliers, represented as individual points beyond the whiskers. These outliers indicate months where specific weather stations recorded significantly lower temperatures compared to the overall distribution. For instance, in January and March, some stations reported average temperatures around 20°F, while in May, June, August, October, November, and December, outliers around 60°F were observed. These anomalies may be due to extreme weather events, data collection errors, or unique microclimatic conditions at specific stations. Further investigation into these outliers can help validate the data and understand the underlying climatic factors.



Boxplot of Monthly Average Precipitation

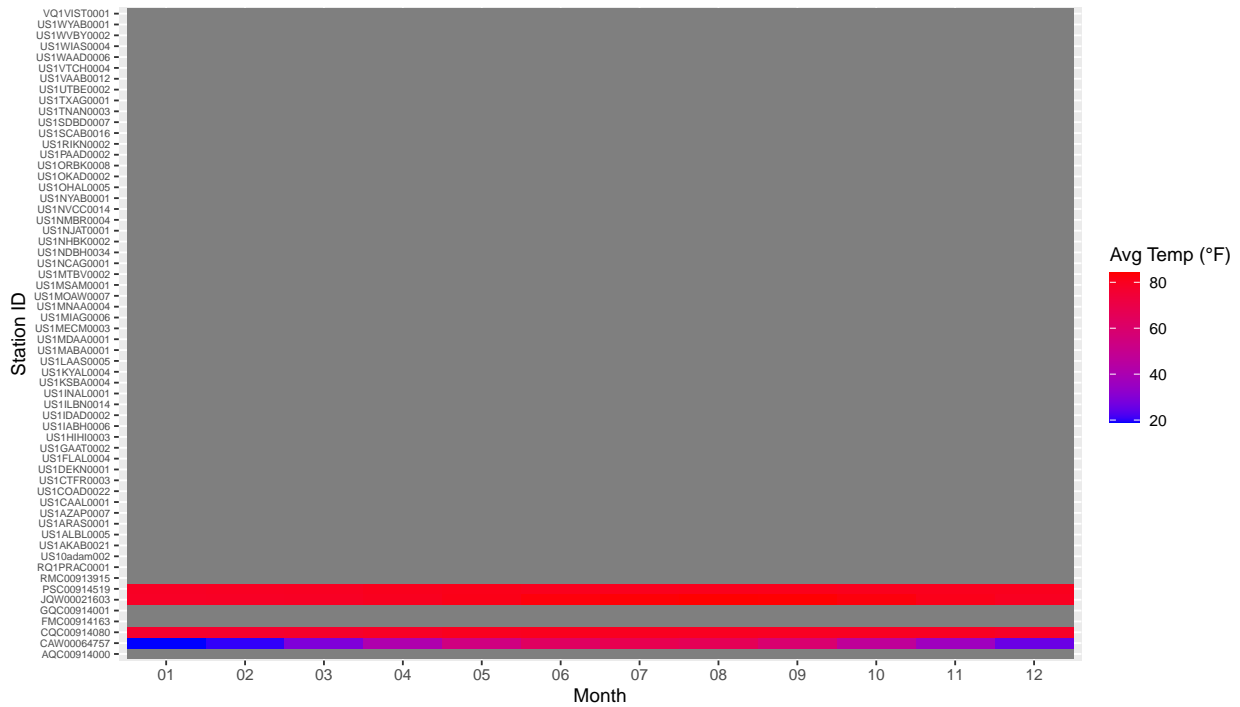
The boxplot of monthly average precipitation shows the distribution of precipitation levels for each month. Outliers highlight months with unusually high or low precipitation, which could be due to extreme weather events or other factors.



Heatmap of Monthly Average Temperature

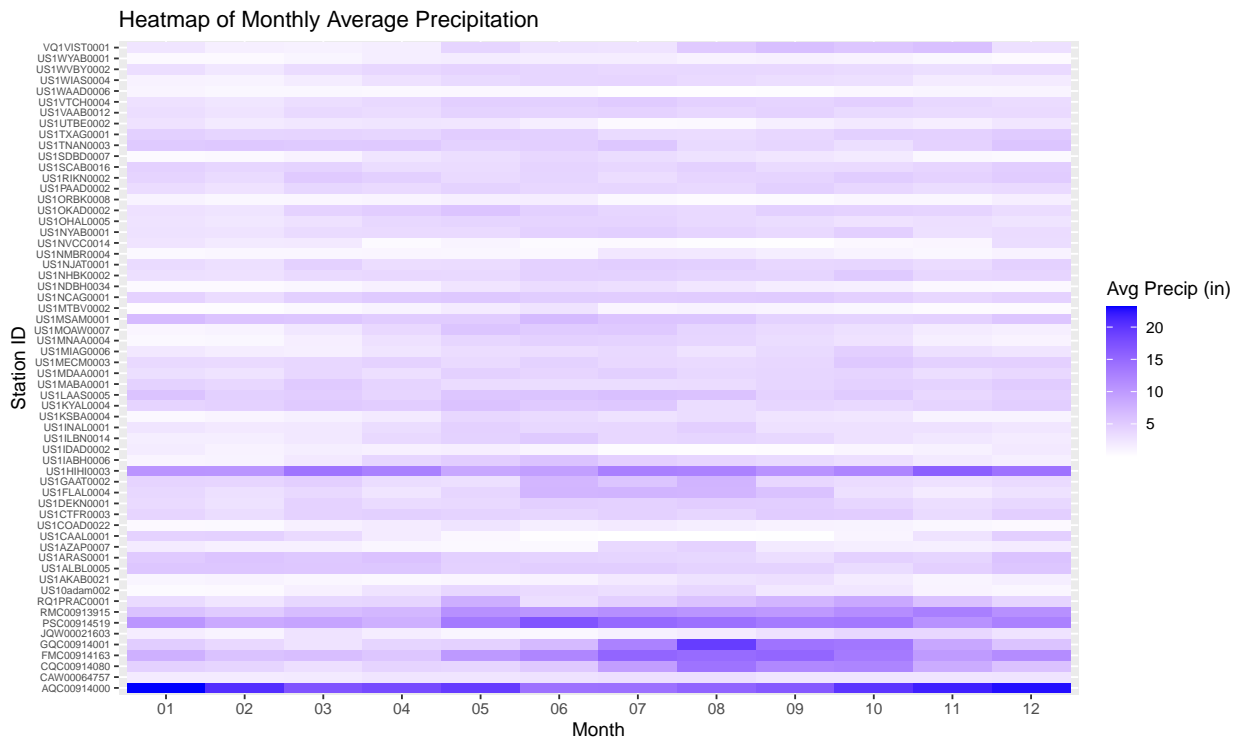
The heatmap of monthly average temperature visualizes the temperature patterns across different stations and months. The color gradient highlights the variations, with warmer colors indicating higher temperatures and cooler colors indicating lower temperatures.

Heatmap of Monthly Average Temperature

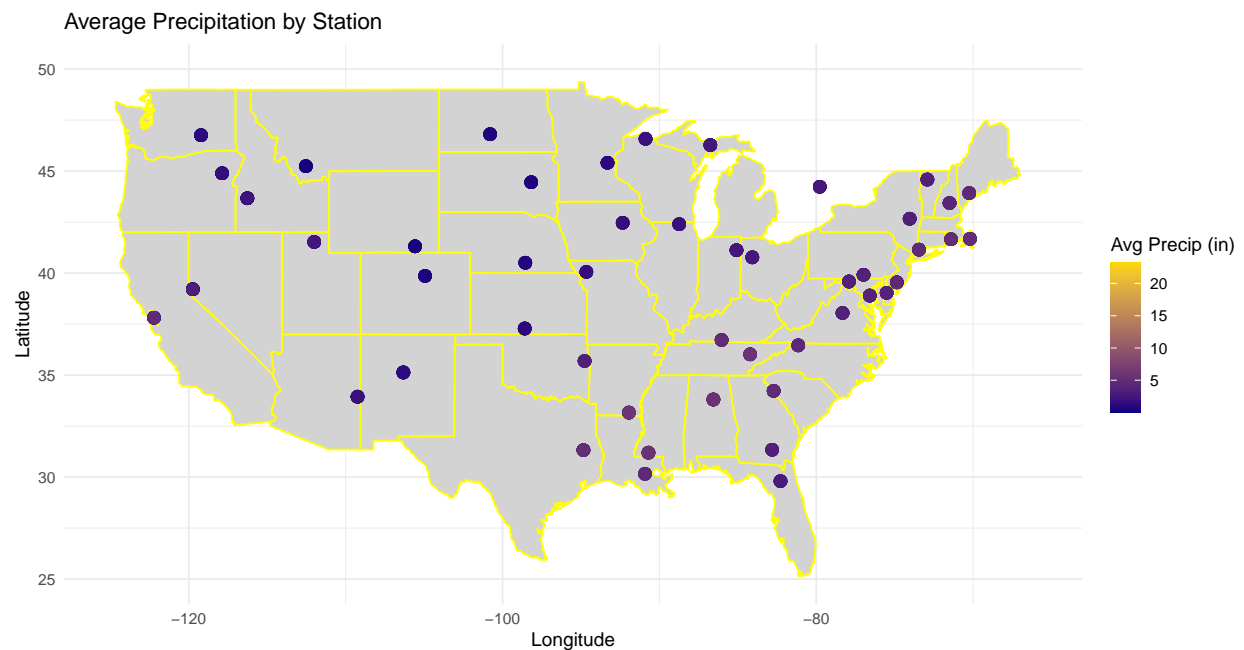
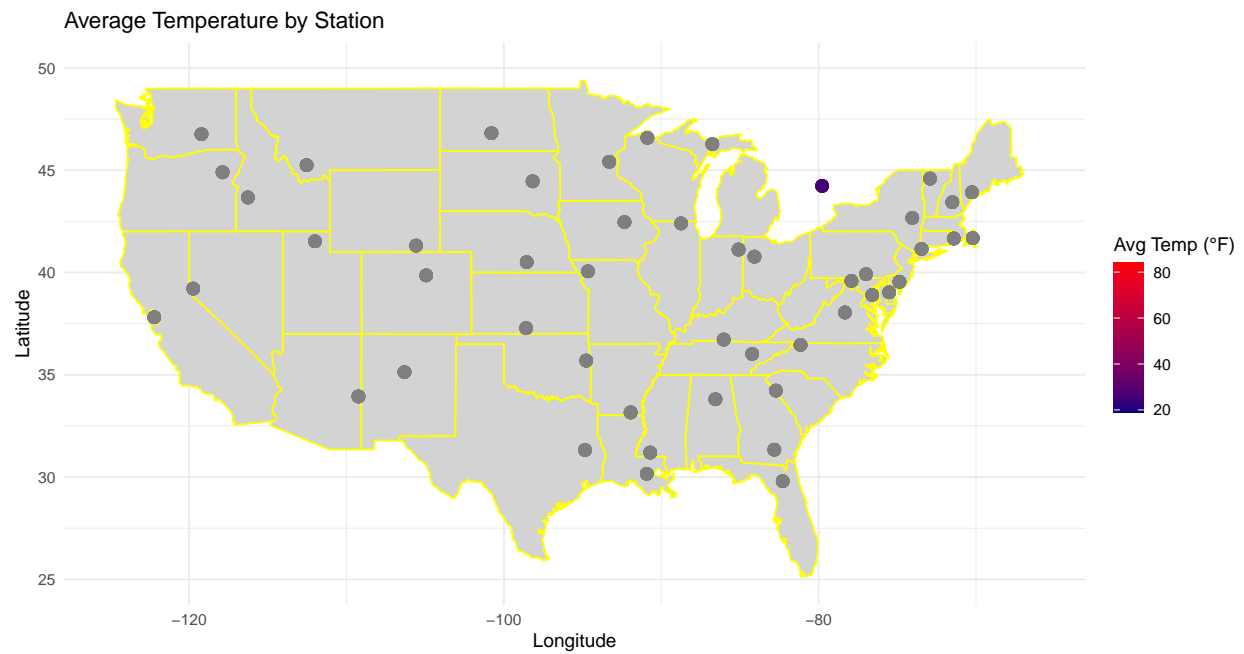


Heatmap of Monthly Average Precipitation

The heatmap of monthly average precipitation displays the precipitation patterns across different stations and months. The color gradient highlights the variations, with darker colors indicating higher precipitation levels.



The following maps show the geographical distribution of average temperature and precipitation for each station. These visualizations help in understanding the spatial patterns in the climate data.



The following histograms show the distribution of temperatures, precipitation across all stations to help understand the frequency.

