# PROPOSAL OF GHCN DATA ANALYSIS

## Yuhan Meng, Shirley Zhang, Chenghui Li, Jitian Zhao

### STAT 605 Final Project

## Data Description

- **Source**:https://www.ncei.noaa.gov/data/global-historical-climatology-network-daily/
- **Data Context**: Daily climate summaries from thousands of weather stations from all over the world.
- **Data Format**: Each file represents years of records collected at one station. Most of records include ten attributes indicating location, temperature, snowfall and precipitation.
- **Attributes of Interest**: Precipitation, Date, Station

## Computation Methods

- Use parallel computation to select data files that meet our requirements.
- Use parallel computation to fit HoltWinters model for our selected data files.
- Select stations that have significant trend in certain attributes and explore their periodicity.

## Questions & Methods

- **Are there any stations indicate significant monthly periodicity of precipitation?**
- **Time Series Analysis, ARCH Effect Analysis.**

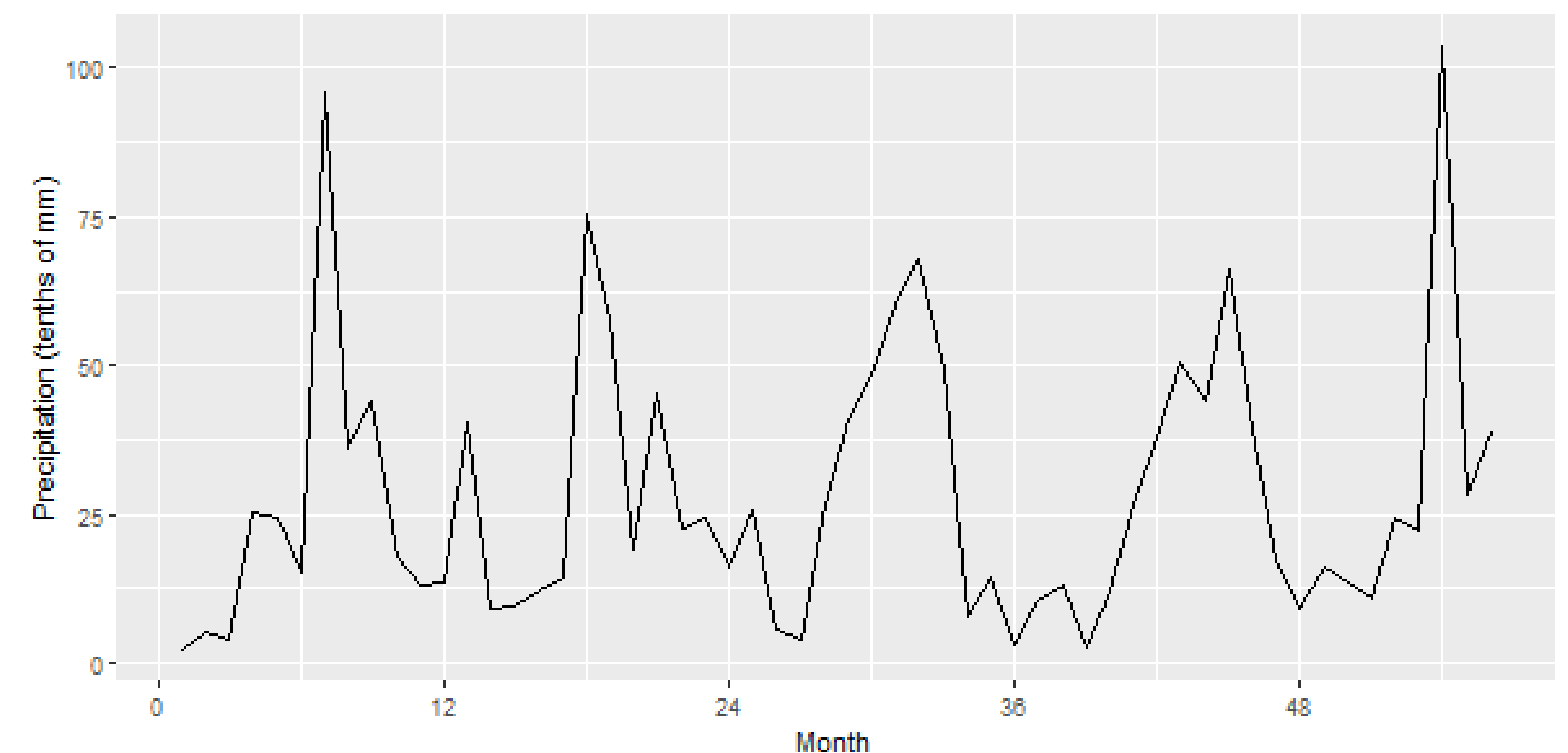## Time Series Plot

**data = read.csv('../data/CA00101C0ME.csv')**
**plot(ts(prcp_monthly))**



Fig. 1: Monthly Average Precipitation