



ANALYSIS OF GLOBAL CLIMATE CHANGE

Xinkai Chen, Siqi Shen, Jiawei Wu, Bi Qing Teng

Introduction

Aim

- To study how temperature has changed over the years

Data

- NOAA Global Historical Climatology Network Daily dataset
(<https://www.kaggle.com/noaa/noaa-global-historical-climatology-network-daily>)

Variables

- station ID
- date
- maximum and minimum temperature
- daily precipitation
- snowfall and snow depth

Methods

- Regression
- Parallel Computation
- Analysis

Regression (Year 1763)

Example of data set:

Station ID	Date	Type of Variables	Observations
ITE00100554	17630101	TMAX	-36
ITE00100554	17630101	TMIN	-50
ITE00100554	17630102	TMAX	-26
ITE00100554	17630102	TMIN	-40
ITE00100554	17630103	TMAX	-9
ITE00100554	17630103	TMIN	-29

Results:

Average intercepts of all sites	Average slopes of all sites
11.90	0.016

Parallel Computation

Year 1763 - 2000

Linear regression for
each site every year

Parallel
computation
CHTC
(238 jobs)

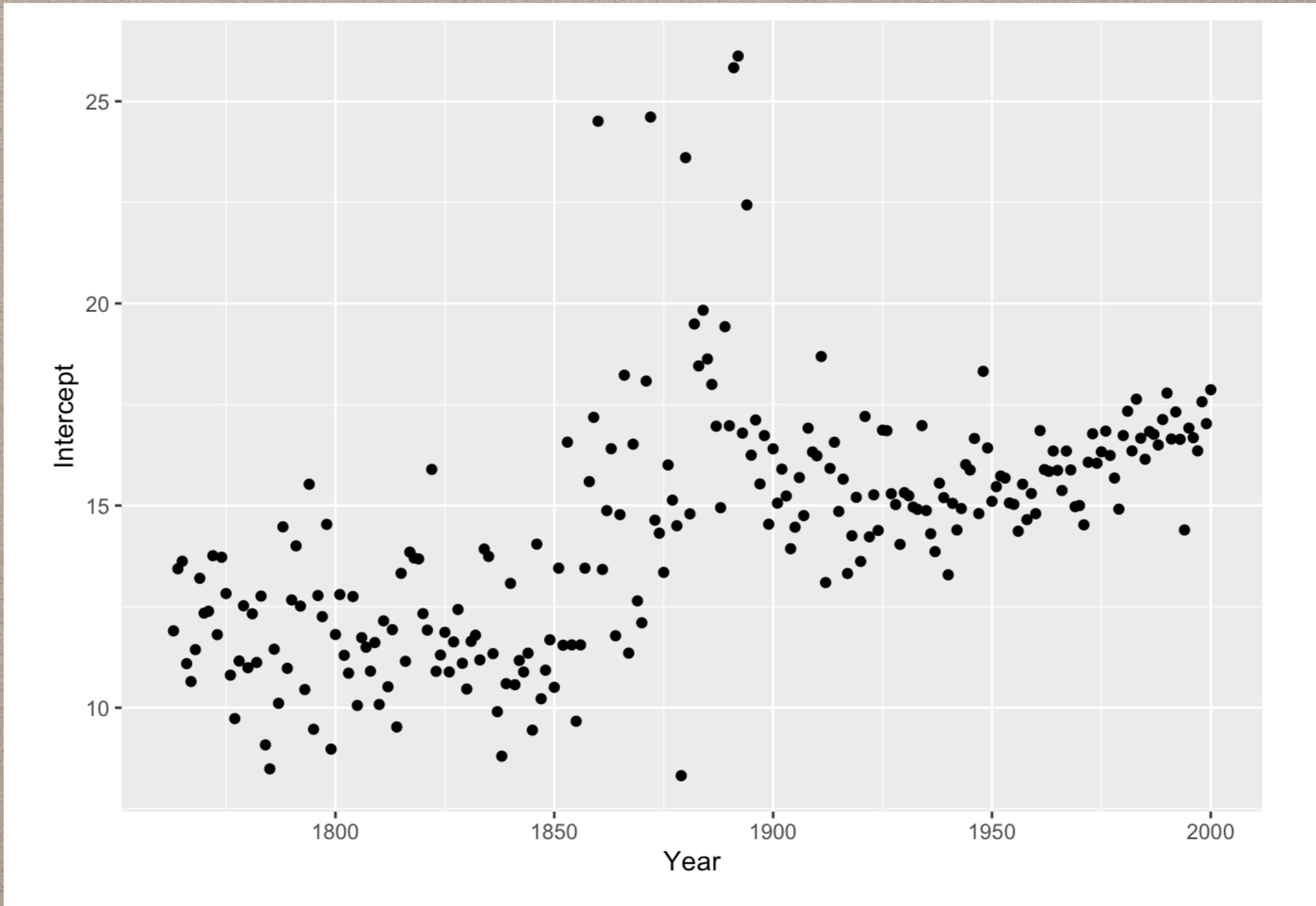
Merge files

238 output files
- Mean intercept
- Mean slope

Parallel Computation

File Names	Jobs #	Memory	Time
1763.csv(23kb)- 1909.csv	147	2 GB	80 min
1910.csv-1937.csv	28	5 GB	120 min
1938.csv- 2000.csv(1GB)	63	10 GB	140 min

Results



Conclusions

The Industrial Revolutions have a harmful impact on our environment

The Third Industrial Revolution is ongoing

Future Work

Split the stations by different areas or countries

Precipitation Analysis



THANK YOU!