CS653
Final Project Report

The Impact of *Climate Change* on Winter Tourism

Members

Data Engineer Role

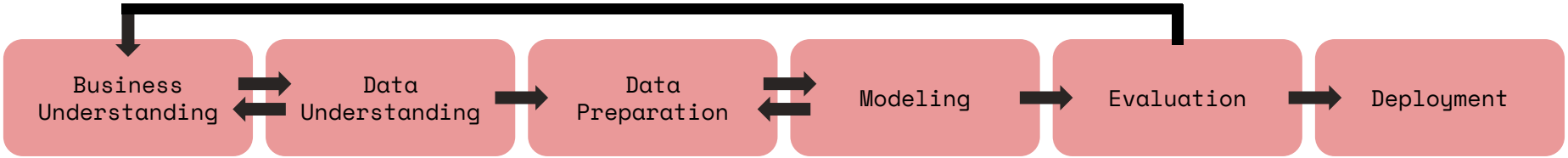
Ratchapol Marmongkol	6509035033
Prerapong Ramunudom	6509035223
Worrapon Thongmook	6509035025

Data Scientists Role

Teevit Lertchaturaporn	6509035082
Kunanon Sukjaruen	6509035207
Rattakorn Keatprakob	6509035165

Our study followed the CRISP-DM methodology (Cross-Industry Standard Process for Data Mining) for the data analysis process.

CRISP-DM



- Introduction
- Objectives

- NOAA (temperature)
- Snow Area

- Data pipeline design

Business Understanding

Introduction

Global warming and climate change are having a significant impact on the world's weather patterns

Temperatures are rising worldwide, and this is leading to changes in precipitation patterns.

It causing changes in the timing of snowfall, which is making it more difficult to plan for winter activities.

Objectives

The purpose of this study is to assess the impact of climate change on the U.S. land temperature and average snow pattern. The study will examine the following:

1. The use of Amazon Web Services (AWS) to utilize big data analytics to tackle the global warming and climate change problem.
2. The impact of the U.S. land temperature on snow coverage area.
3. The analytics of time series data on how the U.S. tourism industry needs to adapt the strategies.

Data Understanding

Data Understanding

(1) EpiNOAA - NOAA U.S. Climate Gridded Dataset (NClimGrid)

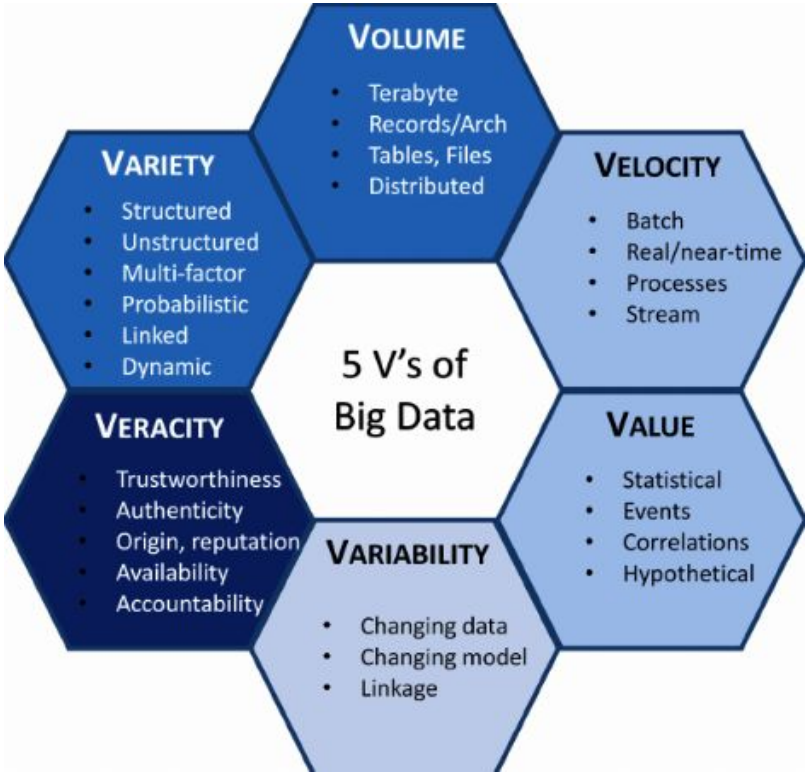
Field	Description and format
date	The date of the weather observation (format : mm/dd/yyyy)
year	The year of the weather observation (format : yyyy)
month	The month of the weather observation (format : mm)
day	The day of the month of the weather (format : dd)
state	The state in which the weather observation was taken (format : XX two abbreviations characters)
county	The sub-area of the state in which the weather observation was taken
region_code	The code which indicate state-county
prcp	The amount of precipitation that fell on the day of the weather observation in millimeters (mm)
tavg	The average temperature for the day of the weather observation in degree celsius (^{O}C)
tmin	The minimum temperature for the day of the weather observation in degree celsius (^{O}C)
tmax	The maximum temperature for the day of the weather observation in degree celsius (^{O}C)

Data Understanding

(2) Monthly Area of Snow Extent

Field	Description and format
Year	The year of the snow cover observation
Month	The month of the snow cover observation
N.Hemisphere	The area of snow cover in the Northern Hemisphere in million square kilometers
Eurasia	The area of snow cover in Eurasia in million square kilometers
N.America	The area of snow cover in North America in million square kilometers
N.America (no Greenland)	The area of snow cover in North America excluding Greenland in million square kilometers

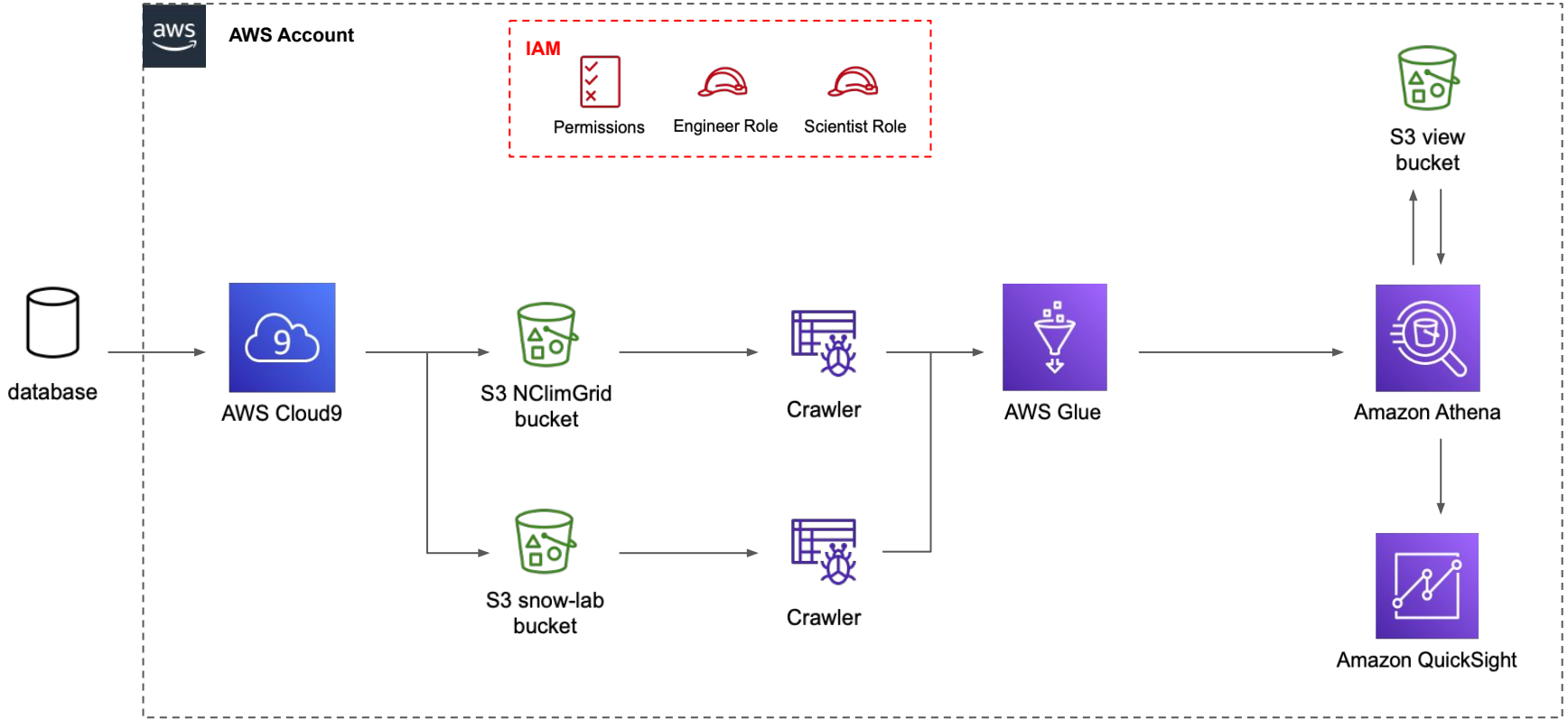
5Vs of Data


VOLUME
Terabyte
Records/Arch
Tables, Files
Distributed
VARIETY
Structured
Unstructured
Multi-factor
Probabilistic
Linked
Dynamic
VELOCITY
Batch
Real/near-time
Processes
Stream
5 V's of Big Data
VERACITY
Trustworthiness
Authenticity
Origin, reputation
Availability
Accountability
VALUE
Statistical
Events
Correlations
Hypothetical
VARIABILITY
Changing data
Changing model
Linkage

Data Pipeline

Design and Implementation

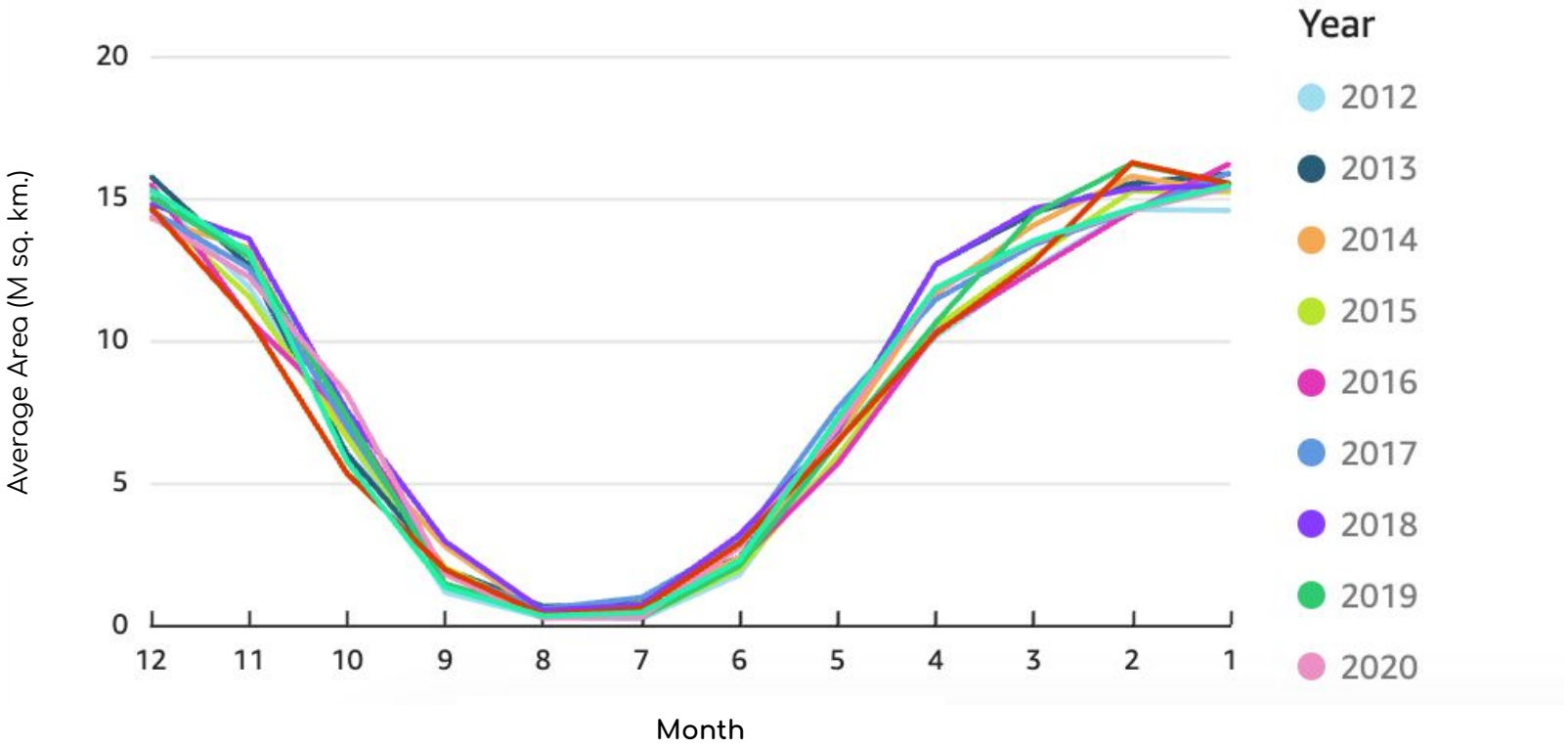
Architecture of Data Pipeline



Result

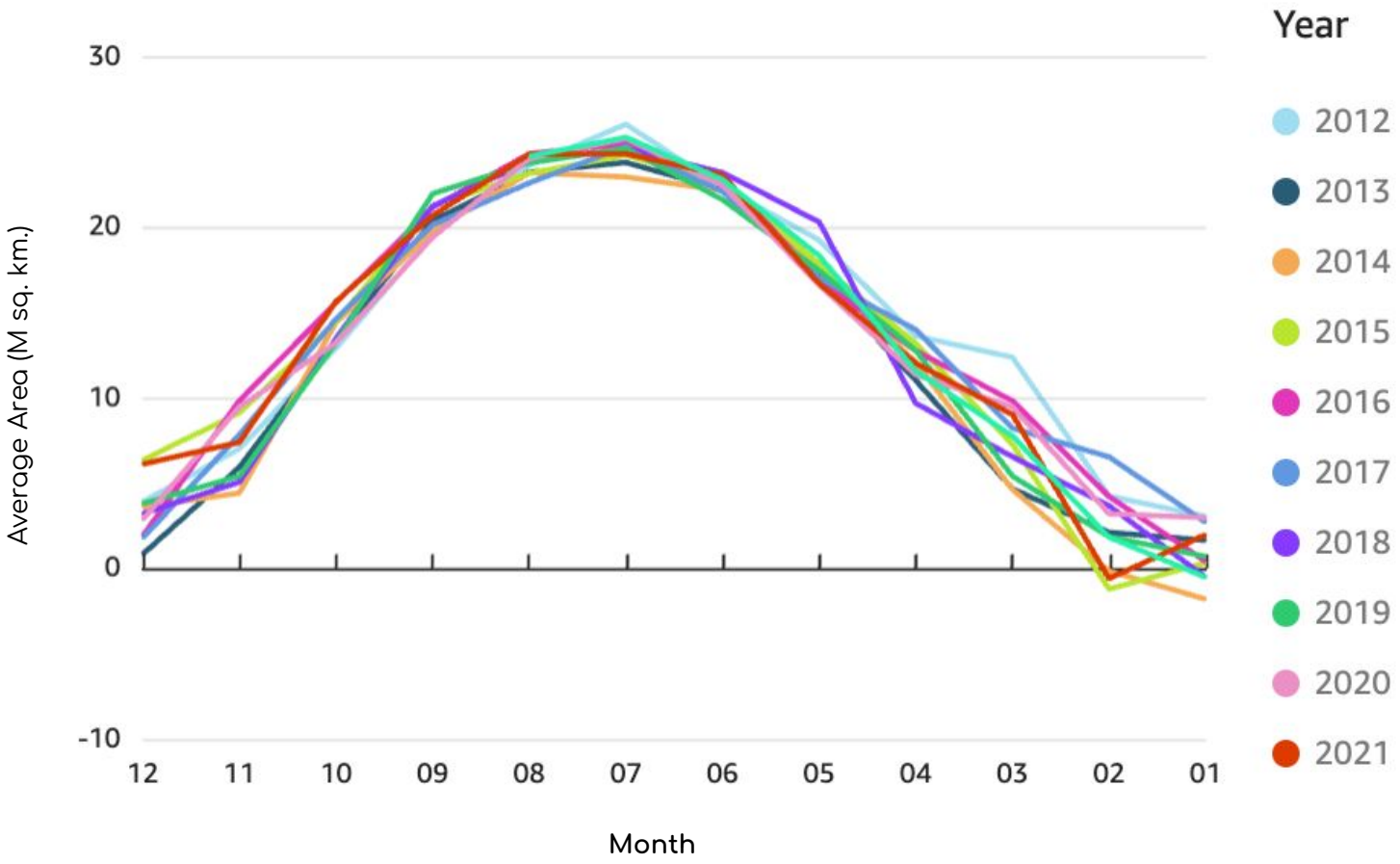
Monthly average U.S. area of snow coverage in each year

2012 - 2020



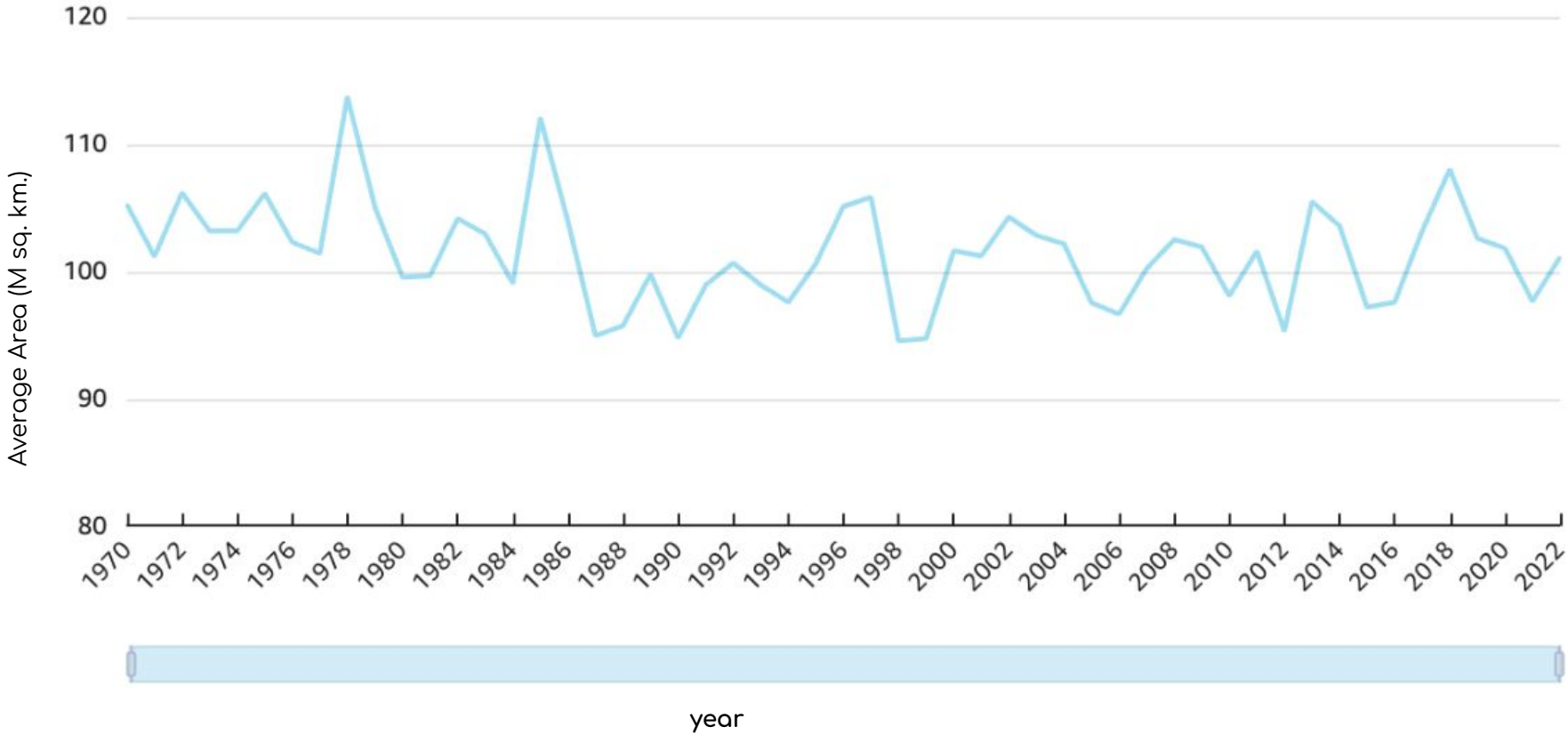
Monthly average U.S. land temperature in each year

2012 - 2021



Area of snow coverage in each year

1970 - 2022



Evaluation and Conclusion