

Exploratory Analysis of GHCN Data with Spark

David Ewing (82171165)

March 24, 2025

Abstract

This report documents the analysis of GHCN climate data using Apache Spark in a Kubernetes-based cloud environment.

1 Introduction

Write your introduction here.

2 Methods

Describe your data processing and analysis steps.

3 Results

Summarise your key findings.

4 Discussion

Reflect on what the results mean, potential limitations, and next steps.