

COMP1204: Data Management

Coursework One: Hurricane Monitoring

John Holman
jh11g20

March 15, 2022

1 Introduction

The aims of this work is to create a bash script that will take a kml file containing storm tracking data and clean it so that it is formatted to be readable and can be passed to other scripts such as to display the data visually. It involves identifying the relevant information sections and removing excess code and formatting to leave only the data itself, then formatting this into a specific arrangement.

2 Create CSV Script

```
#!/bin/bash

#sets variables for the file arguments to read and write
filename=$1
outputname=$2

#takes file contents, greps only the relevant data area then trims all the extra metadata and kml formatting as well as the titles declaring what each value is. Then adds newlines at the end of knots so that each line of data is defined
content=$(cat $filename | grep '<tr><td>' | sed 's#<tr><td>##' | sed 's#</td></tr>##' | sed 's#</table>##')

#outputs to the file the first line of the collumn names
echo "Timestamp,Latitude,Longitude,MinSeaLevelPressure,MaxIntensity" > $outputname

#writes content to a text file so that the while loop functions as it needs a file to work with
echo "$content" > MidWayFile.txt

#while loop reads each line of content and writes it to the file, then pipes reformat the commas in the right places and removes the spaces at the end of the commas, finally output is written to the output file without overwriting the collumn names from before
input="MidwayFile.txt"
while IFS= read -r line
do
    echo "$line"
done < "$input" | awk '{$4=$4","}' | awk '{$10=$10","}' | awk '{$8=$8","}' | sed 's#, #,#g' >> $output
```

3 Storm Plots

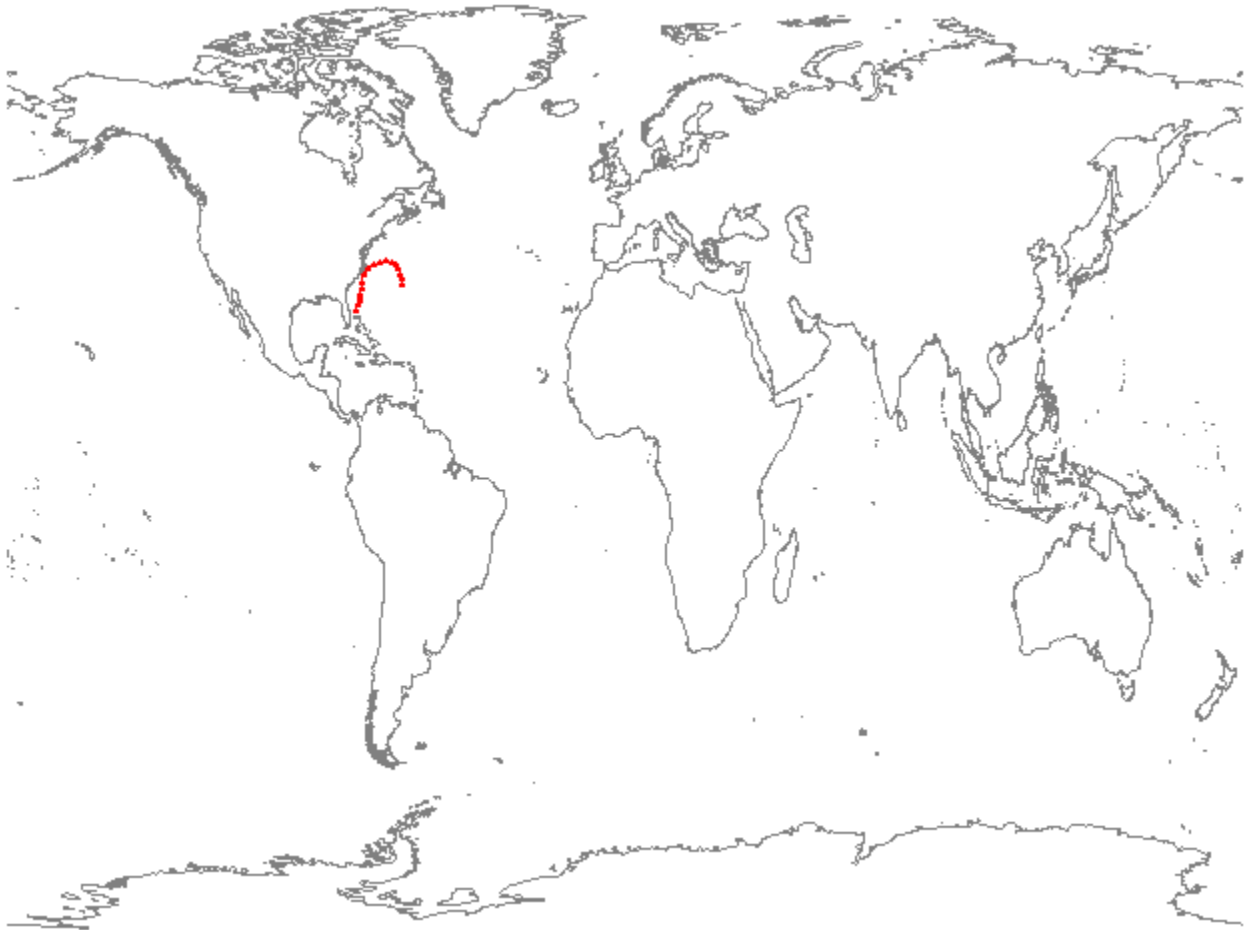


Figure 1: al012020 storm plot

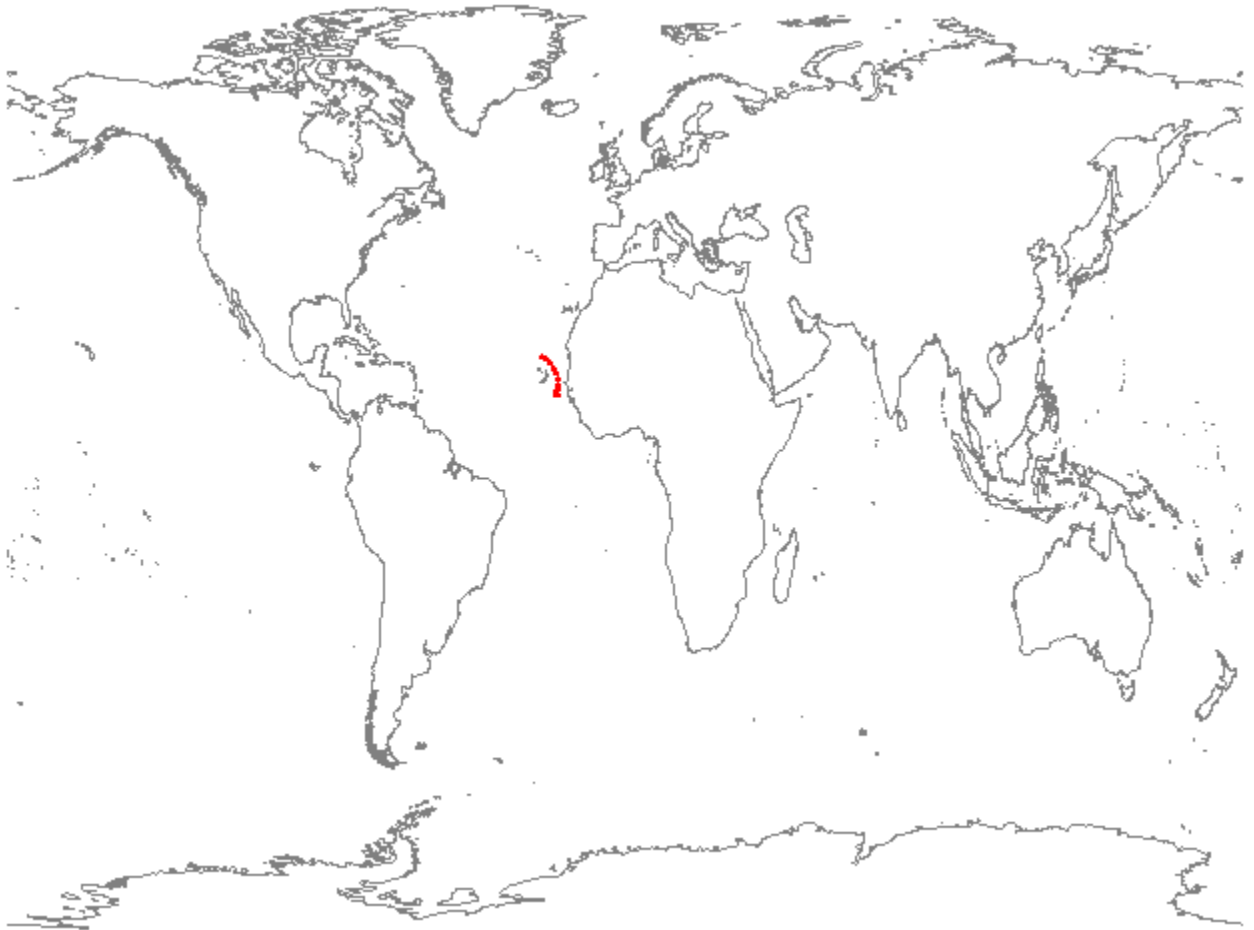


Figure 2: al102020 storm plot

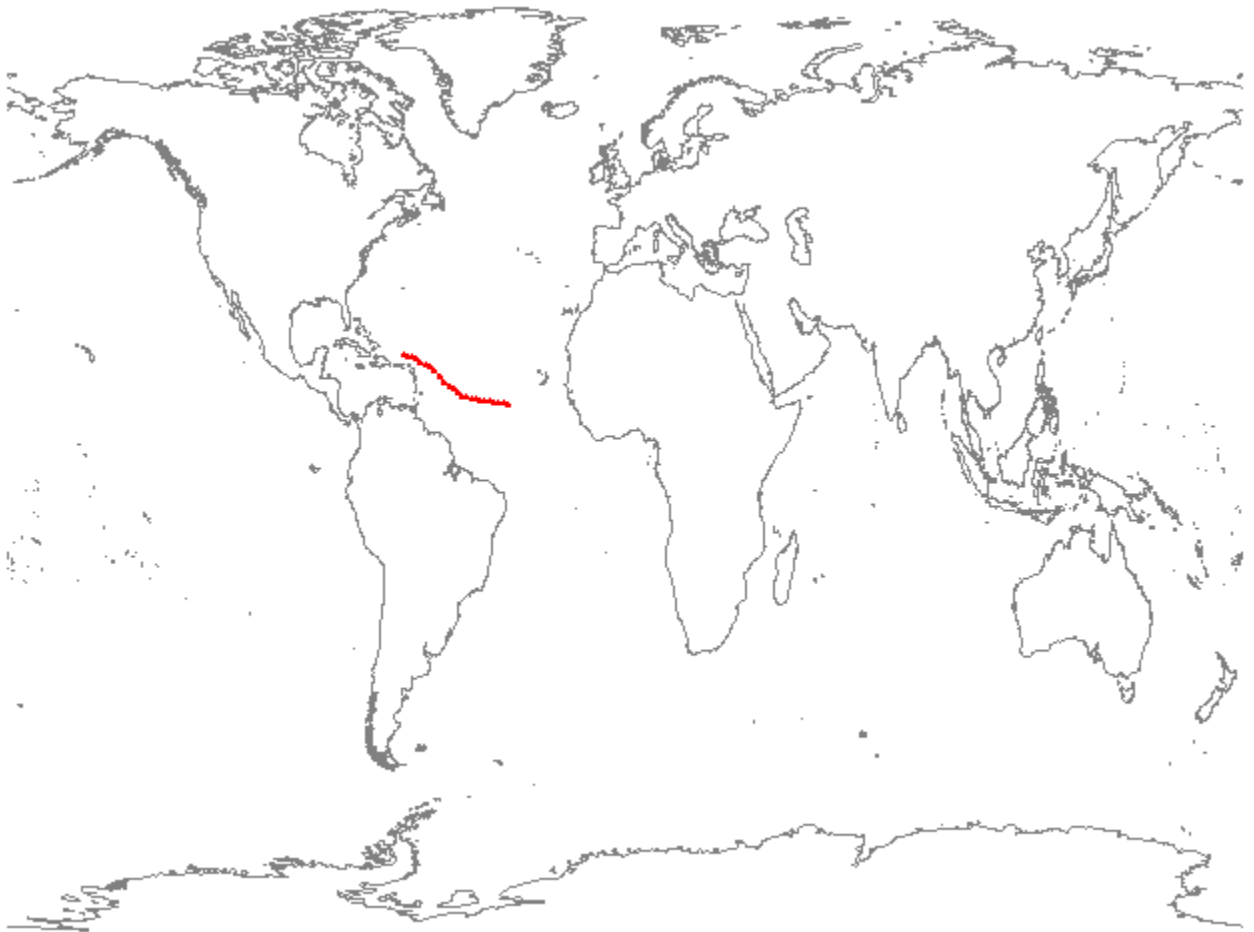


Figure 3: all12020 storm plot

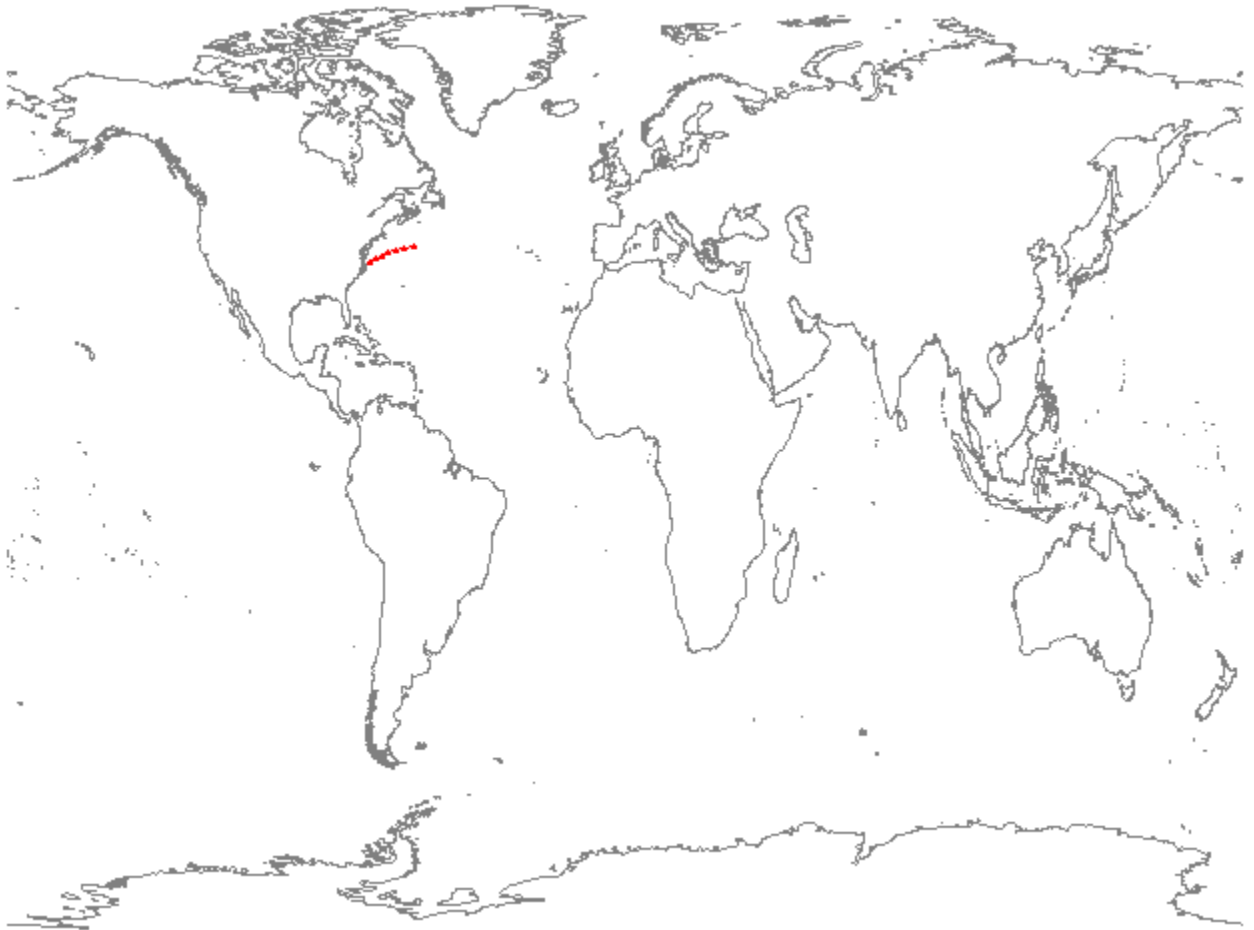


Figure 4: al122020 storm plot

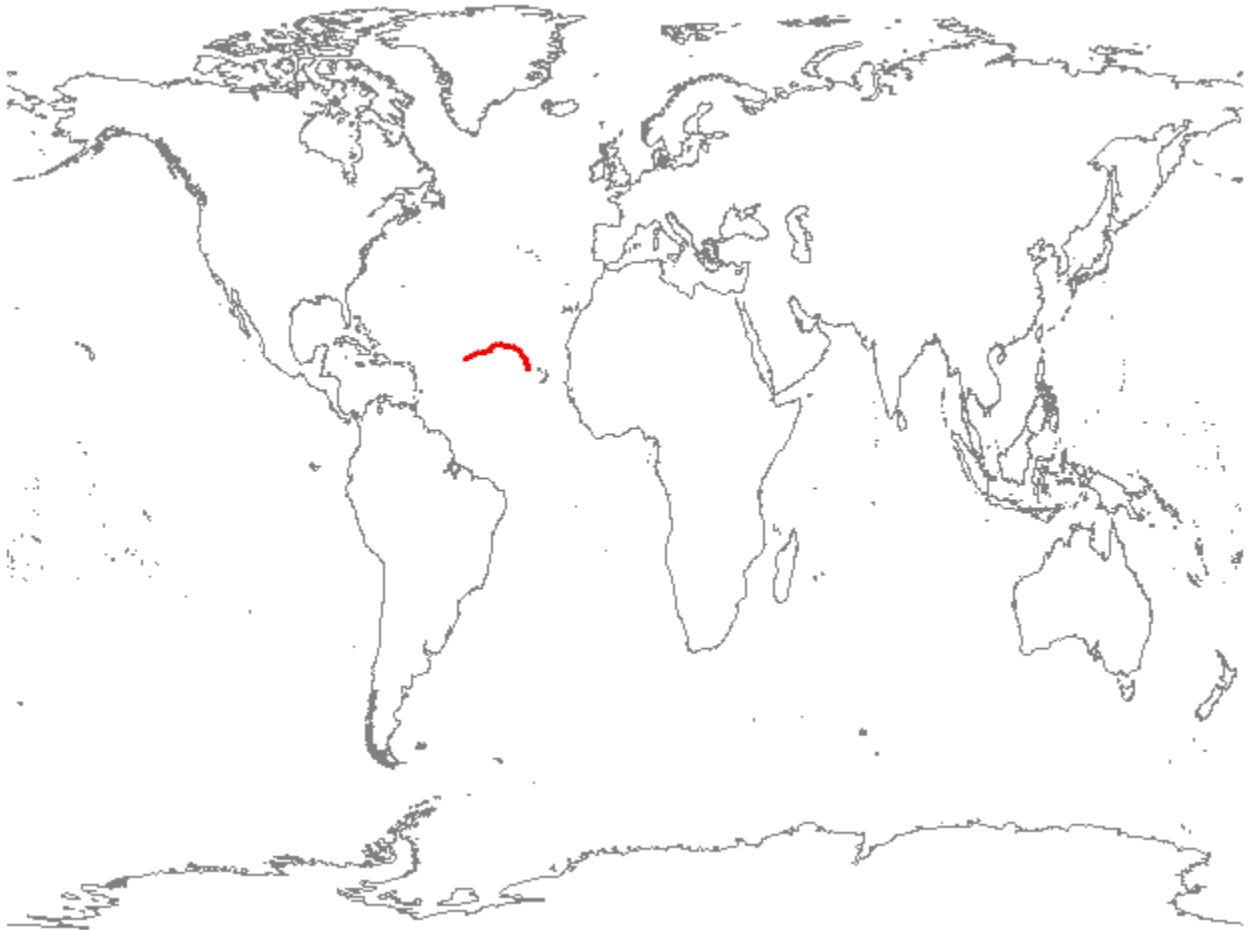


Figure 5: al212020 storm plot