**Assignment 5 – Oisin Mc Laughin 22441106**

Code:

%Part 1

flights = readtable('Flights.csv');

num\_records = height(flights);

fprintf('Number of records: %d\n', num\_records);

%Part 2

flights.origin = string(flights.origin);

flights.dest = string(flights.dest);

%Part 3

dep\_delay = flights.dep\_delay;

minDelay = min(dep\_delay);

medianDelay = median(dep\_delay, 'omitnan');

maxDelay = max(dep\_delay);

numMissing = sum(ismissing(dep\_delay));

fprintf('Min: %d\n', minDelay);

fprintf('Median: %d\n', medianDelay);

fprintf('Max: %d\n', maxDelay);

fprintf('NumMissing: %d\n', numMissing);

%Part 4

filtered = flights(~isnan(dep\_delay), :);

fprintf("Filtered size: %d\n", size(filtered));

%Part 5

fprintf('Flights vs Flights Filtered: %d\n', (height(flights) - height(filtered)));

%Part 6

final = flights(dep\_delay <= 120, :);

fprintf('Final size: %d\n', (height(final)));

%Part 7

Months = unique(flights.Month);

avgDelays = zeros(length(Months), 1);

for i = 1:length(Months)

currentMonthFlights = flights(flights.Month == Months(i), :);

avgDelays(i) = mean(currentMonthFlights.dep\_delay, 'omitnan');

end

res1 = table(Months, avgDelays, 'VariableNames', {'Month', 'AvrDelayMonth'});

disp(res1);

plot(Months, avgDelays, '-o');

title('Average Delay By Month');

xlabel('Month');

ylabel('Average Delay (minutes)');

%Part 8

hours = unique(flights.hour);

avgDelayPerHour = zeros(length(hours), 1);

figure;

for i = 1:length(hours)

currentHourFlights = flights(flights.hour == hours(i), :);

avgDelayPerHour(i) = mean(currentHourFlights.dep\_delay, 'omitnan');

end

res2 = table(hours, avgDelayPerHour, 'VariableNames', {'Hour', 'AvrDelayHour'});

disp(res2);

plot(hours, avgDelayPerHour, '-o');

title('Average Delay By Hour of the Day');

xlabel('Hour');

ylabel('Average Delay (minutes)');

%Part 9

Origins = unique(flights.origin);

Months = unique(flights.Month);

avgDelaysMonthOrigin = zeros(length(Months), length(Origins));

for i = 1:length(Months)

for j = 1:length(Origins)

currentFlights = flights(flights.Month == Months(i) & flights.origin == Origins(j), :);

avgDelaysMonthOrigin(i, j) = mean(currentFlights.dep\_delay, 'omitnan');

end

end

MonthCol = repmat(Months, length(Origins), 1);

OriginCol = repelem(Origins, length(Months));

avgDelaysCol = avgDelaysMonthOrigin(:);

res3 = table(MonthCol, OriginCol, avgDelaysCol, 'VariableNames', {'Month', 'Origin', 'AvrDelayMonthOrigin'});

disp(res3);

figure;

for i = 1:length(Origins)

subplot(length(Origins), 1, i);

plot(Months, avgDelaysMonthOrigin(:, i), '-o');

title(sprintf('Average Delay By Month: %s', Origins(i)));

xlabel('Month');

ylabel('Average Delay (minutes)');

xlim([1 12]);

end

Output:

flightsFunc

Number of records: 336776  
Min: -43  
Median: -2  
Max: 1301  
NumMissing: 8255  
Filtered size: 328521  
Filtered size: 9  
Flights vs Flights Filtered: 8255  
Final size: 318798  
 **Month**  **AvrDelayMonth**  
 **\_\_\_\_\_**  **\_\_\_\_\_\_\_\_\_\_\_\_\_**  
  
 1 10.037   
 2 10.817   
 3 13.227   
 4 13.938   
 5 12.987   
 6 20.846   
 7 21.728   
 8 12.611   
 9 6.7225   
 10 6.244   
 11 5.4354   
 12 16.577

**Hour**  **AvrDelayHour**  
 **\_\_\_\_**  **\_\_\_\_\_\_\_\_\_\_\_\_**  
  
 1 NaN   
 5 0.68776   
 6 1.6428   
 7 1.9141   
 8 4.1279   
 9 4.5837   
 10 6.4983   
 11 7.1917   
 12 8.6148   
 13 11.438   
 14 13.819   
 15 16.895   
 16 18.757   
 17 21.101   
 18 21.11   
 19 24.785   
 20 24.304   
 21 24.196   
 22 18.791   
 23 14.017

clear

flightsFunc

Number of records: 336776  
Min: -43  
Median: -2  
Max: 1301  
NumMissing: 8255  
Filtered size: 328521  
Filtered size: 9  
Flights vs Flights Filtered: 8255  
Final size: 318798  
 **Month**  **AvrDelayMonth**  
 **\_\_\_\_\_**  **\_\_\_\_\_\_\_\_\_\_\_\_\_**  
  
 1 10.037   
 2 10.817   
 3 13.227   
 4 13.938   
 5 12.987   
 6 20.846   
 7 21.728   
 8 12.611   
 9 6.7225   
 10 6.244   
 11 5.4354   
 12 16.577

**Hour**  **AvrDelayHour**  
 **\_\_\_\_**  **\_\_\_\_\_\_\_\_\_\_\_\_**  
  
 1 NaN   
 5 0.68776   
 6 1.6428   
 7 1.9141   
 8 4.1279   
 9 4.5837   
 10 6.4983   
 11 7.1917   
 12 8.6148   
 13 11.438   
 14 13.819   
 15 16.895   
 16 18.757   
 17 21.101   
 18 21.11   
 19 24.785   
 20 24.304   
 21 24.196   
 22 18.791   
 23 14.017

**Month**  **Origin**  **AvrDelayMonthOrigin**  
 **\_\_\_\_\_**  **\_\_\_\_\_\_**  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**  
  
 1 "EWR" 14.906   
 2 "EWR" 13.067   
 3 "EWR" 18.102   
 4 "EWR" 17.4   
 5 "EWR" 15.366   
 6 "EWR" 22.471   
 7 "EWR" 22.035   
 8 "EWR" 13.494   
 9 "EWR" 7.291   
 10 "EWR" 8.6364   
 11 "EWR" 6.7238   
 12 "EWR" 21.027   
 1 "JFK" 8.6158   
 2 "JFK" 11.791   
 3 "JFK" 10.722   
 4 "JFK" 12.249   
 5 "JFK" 12.519   
 6 "JFK" 20.5   
 7 "JFK" 23.769   
 8 "JFK" 12.914   
 9 "JFK" 6.6358   
 10 "JFK" 4.5926   
 11 "JFK" 4.6785   
 12 "JFK" 14.788   
 1 "LGA" 5.6416   
 2 "LGA" 6.9616   
 3 "LGA" 10.232   
 4 "LGA" 11.509   
 5 "LGA" 10.631   
 6 "LGA" 19.297   
 7 "LGA" 18.995   
 8 "LGA" 11.244   
 9 "LGA" 6.2074   
 10 "LGA" 5.3091   
 11 "LGA" 4.7671   
 12 "LGA" 13.589

Graphs:

A graph with blue lines

Description automatically generated

A graph with a line

Description automatically generated

A graph of a number of months

Description automatically generated with medium confidence