STAT9006 Sample Assessment 01 Dr Seán Lacey

The Environmental Protection Agency (EPA) is concerned about the air quality in regions of the southwest, where large industrial estates are located. Ozone readings in parts per million (ppm) were taken at noon at Shannon Airport on 20 consecutive days and the results were recorded as follows

11	13	18	3	15	9	20	15	17	13
17	13	12	24	7	5	17	10	8	14

Ozone readings (ppm) were also taken at noon at Cork Airport on the same 20 consecutive days and those results were recorded as follows:

11	10	20	4	12	8	16	14	13	15
23	11	16	22	6	12	17	7	12	19

As an employee of the EPA, you are required to compare the readings in both locations. The Shannon airport authority claims that the air quality in Shannon is superior to that in Cork.

Requirements: In your assignment, you are expected to investigate whether or not Shannon airport authority's claim is accurate using R to compile the appropriate tables/graphs etc. You are expected to **explain all concepts and procedures** used in the analysis of the data - i.e., all the **descriptive statistics should be fully explained!**

Breakdown of marks (130 marks)

- [Workshop 02] 20 marks for setting up the dataframe in R correctly
- [Workshop 01/02] 20 marks for a clear introduction to the report with reference to the approach been adopted:
  - 15 marks for stating the **type of data** presented, type of **study**, **sample size** and the main **questions to be answered** in the assessment;
  - 5 marks for summaries on the **approach been undertaken** to answer the questions posed - e.g., Section 2 of the report will summarise the data for each group into frequency distribution tables etc.
- [Workshop 03] 20 marks for the correct use and explanation of **frequency** distribution and **relative** frequency distribution tables (**set the class width to 5, with 0 as the start of the first interval**) and **relative frequency** distribution tables:
  - 2.5 marks for each table presented (4 tables/rows to be presented);
  - 10 marks for **three** correct comment/observation on symmetry/outliers/comparisons.
- [Workshop 02] 20 marks for the correct use and explanation of **boxplots**:
  - 5 marks for each graph presented (2 boxplots within the one graph);
  - 10 marks for **three** correct comment/observation on symmetry/outliers/comparisons.

- [Workshop 02] 20 marks for the test of **normality** and interpretation of the result.
  - 5 marks for presentation of results (2 results to be presented);
  - 10 marks for interpretation of results (2 results to be interpreted);
  - 5 marks for clearly outlining the **approach that you would adopt** if the variables were normally distributed and not normally distributed - i.e., state the descriptive statistics that you would adopt for both scenarios.
- [Workshop 02] 20 marks for the correct presentation and explanation of measures of centrality (**mean, median**), variance (**standard deviation, quartiles**) and **min/max** values.
  - 2 marks for the correct presentation on each pair of statistics (i.e., two means, two medians etc.);
  - 2 marks for a correct comment on each **comparison**.
- [Workshop 02/03] 10 marks for a conclusion on whether or not Shannon airport authority's claim is accurate. Justify your suggestion.