Data Visualization Coursework (SET09120)

Davide Pollicino[1]

School of Computing, Edinburgh Napier University, Edinburgh

40401270@live.napier.ac.uk

Abstract. In this paper, we will explore, analyze and discuss the data visualization of a given dataset. The document will be segmented in three different sections: analyzing the outliers found in the dataset, the relationships in the data and analyzing a data visualization example, provided by the coursework statement.

1. Outliers Found

1.1 Outlier 1

The **first outlier** (see Fig. 1), has been discovered comparing the total scores received by each group of professionals (beginner, experienced, professional, very experienced), in accordance with the Wind Direction.

We can notice that all the first three categories of athletes listed before scored almost the same scores for all the possible wind directions, while just the Very experienced group, has been able to score a higher score, independently from the wind direction.

1.1 Outlier 2

The second outlier (see Fig 2), has been discovered using a **histogram**. It clearly shows that among all the records of the dataset, there is just an anomaly presence of records, where the Distance is between 145 and 160.

2. Relationship Found

2.1 Relationship 1

The **first relationship (**see Fig. 3), has been found comparing the offset with the Score.1, in accordance with the Wind.Direction. The data visualization clearly shows that independently from the wind direction, when the offset tends to between around 38 and 41, we will have an extremely high score1.

2.2 Relationship 2

The **second relationship** (see Fig. 4), has been found comparing the relationship between the Score 2 and the age, in accordance to the Wind Direction. We can see that:

When the Wind direction is **North and South**, independently from the age, the Score.2 tends to be very similar.
When the Wind direction is **West**, the Score.2 is **indirectly proportional** to the age.
When the Wind direction is **East,** the Score.2 is **directly proportional** to the athlete's age.

2.3 Relationship 3

The **third pattern** (see Fig. 5), is a clear relationship between the angle and the distance. One the angle of the lunch is 45°, we reach the distance peak.

2.4 Relationship 4

The **fourth relationship** (see Fig. 6), has been found to analyze the offset in accordance with the Wind Direction and grouping by professionalism levels.

The data visualization shows that, independently from the professionalism level and the wind direction, the offset of any professional category in any wind direction are very similar.

2.4 Relationship 5

The **fifth relationship** (see Fig.7), this relationship has been found comparing Gender and offset and grouping by professionalism level.

The data visualization clearly shows data independently from professionalism level, all the males tend to have a very homogenous offset, while for the females, the offset is indirectly proportional to their experience.

3. Visualization Evaluation

In this section, we will review a data visualization example (See Fig.9). We can clearly see that:

- The female icon used creates an important **visual noise**, having different icons sometimes piles up between them.
- Use regulars' bars of bar charts instead; the girl icon will allow anyone to visualize and compare the heights of the bar in an easier and more readable manner.
- The **color palette** used, could lead any reader to actually think there is any correlation between the pairs of countries: Latvia-S.Africa, Australia-Peru, Scotland-India.

An easier, more readable and clearer visualization example that could be used instead of the **Fig. 9,** is showed by the **Fig. 10**

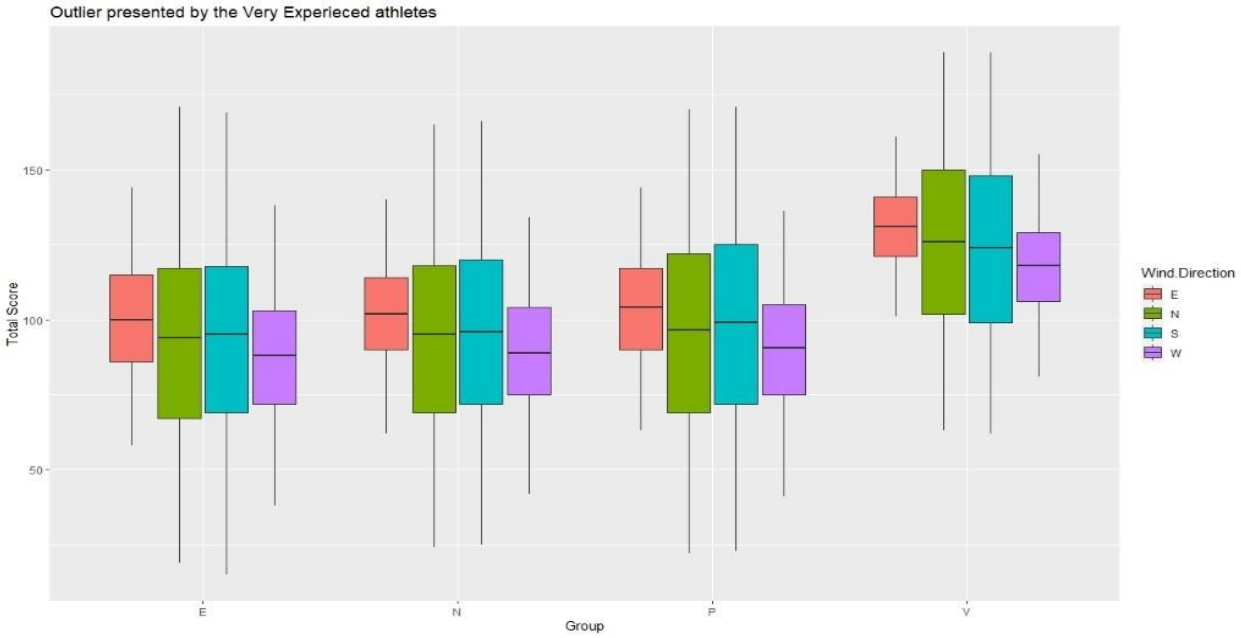

Fig. 1 Relationship between the Age and the Score.2 in according to the Wind Direction

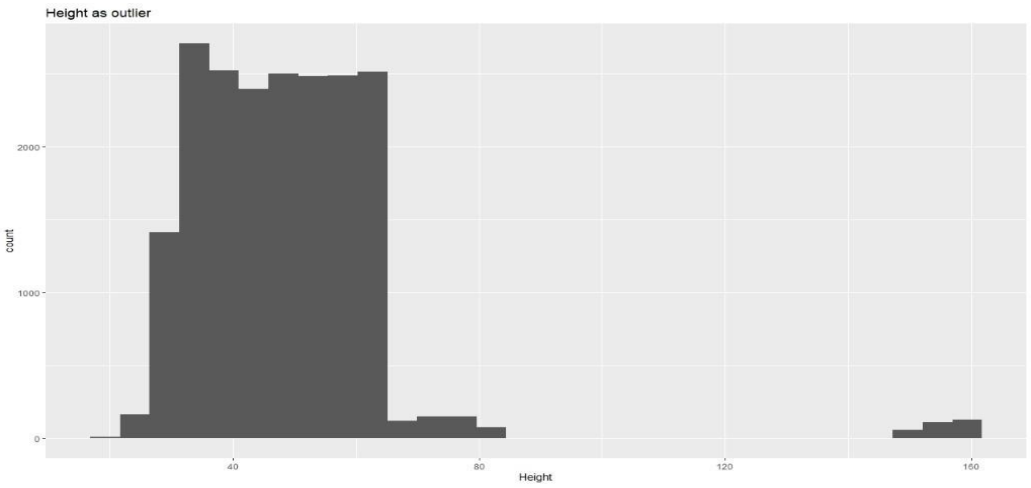

Fig. 2 height As Outlier;

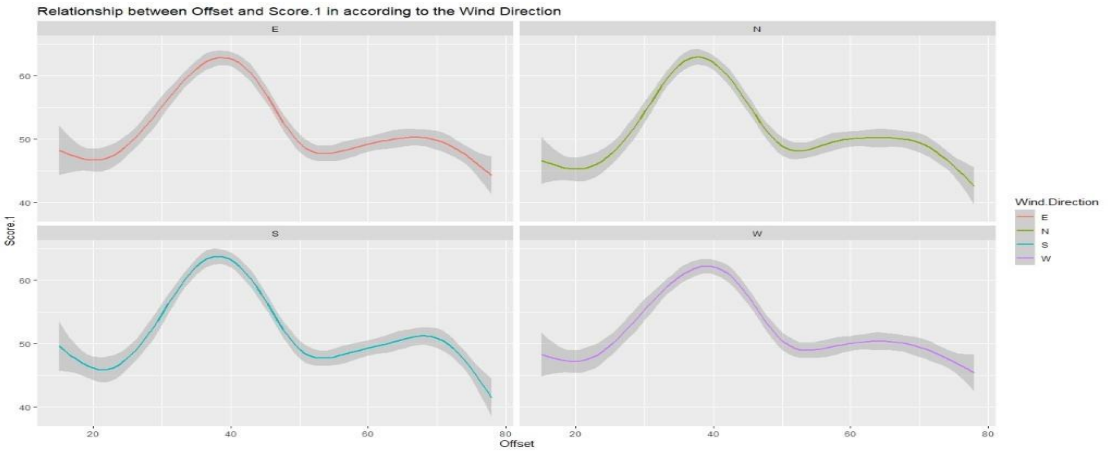

Fig 3: Relationship between Offset and Score.1

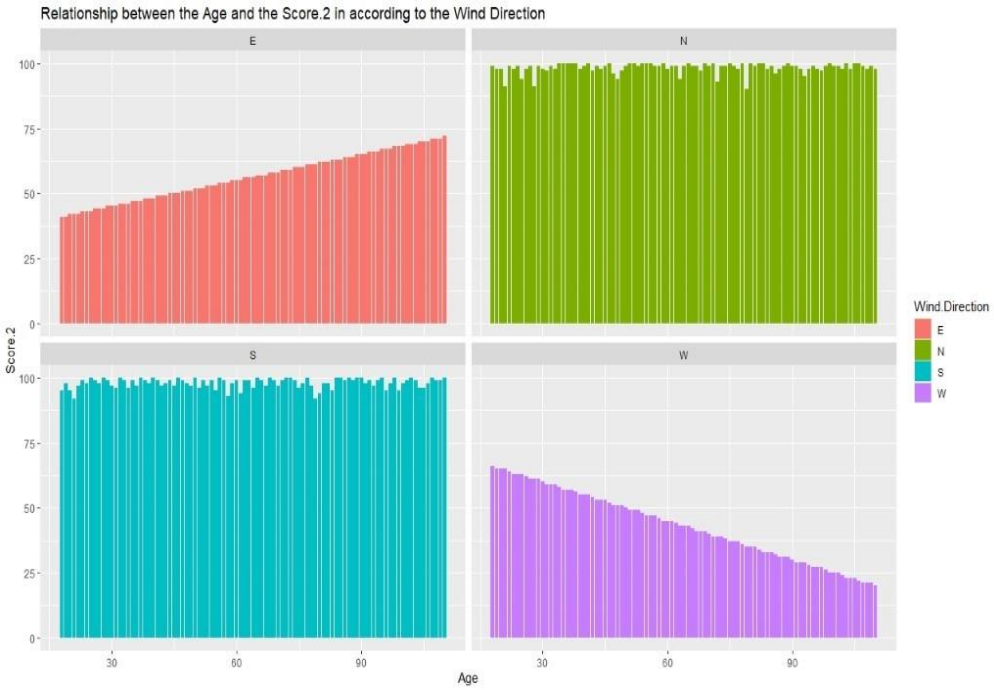

Fig. 4 Relationship between the Age and the Score.2 in according to the Wind Direction

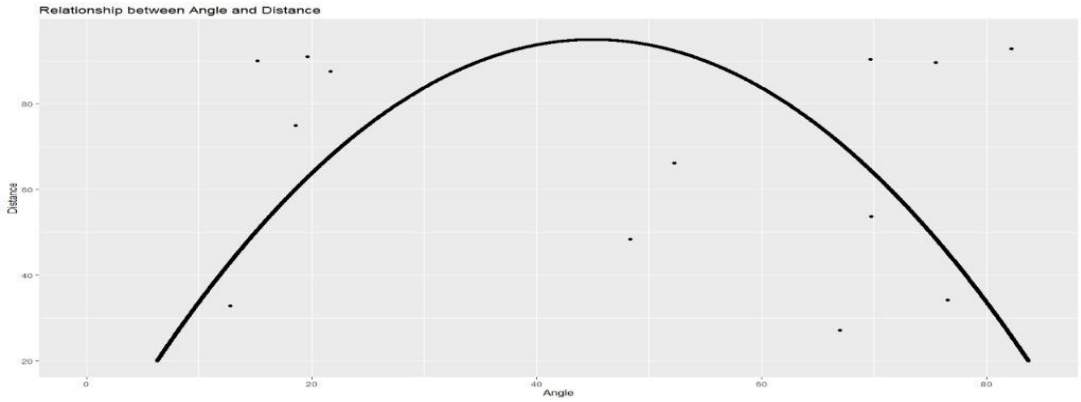

Fig. 5 Relationship between an Angle and the Distance;

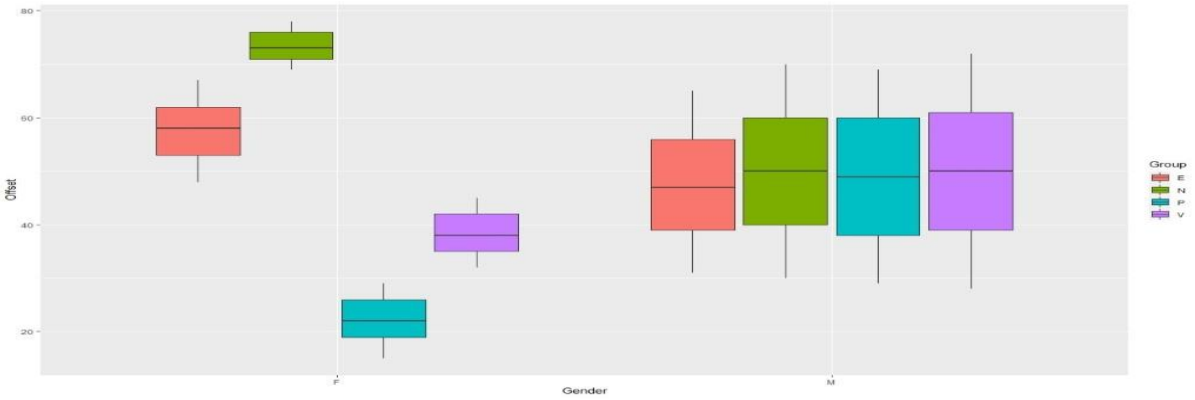

Fig. 7 Relationship between the gender and the Offset in according to the professionalism level

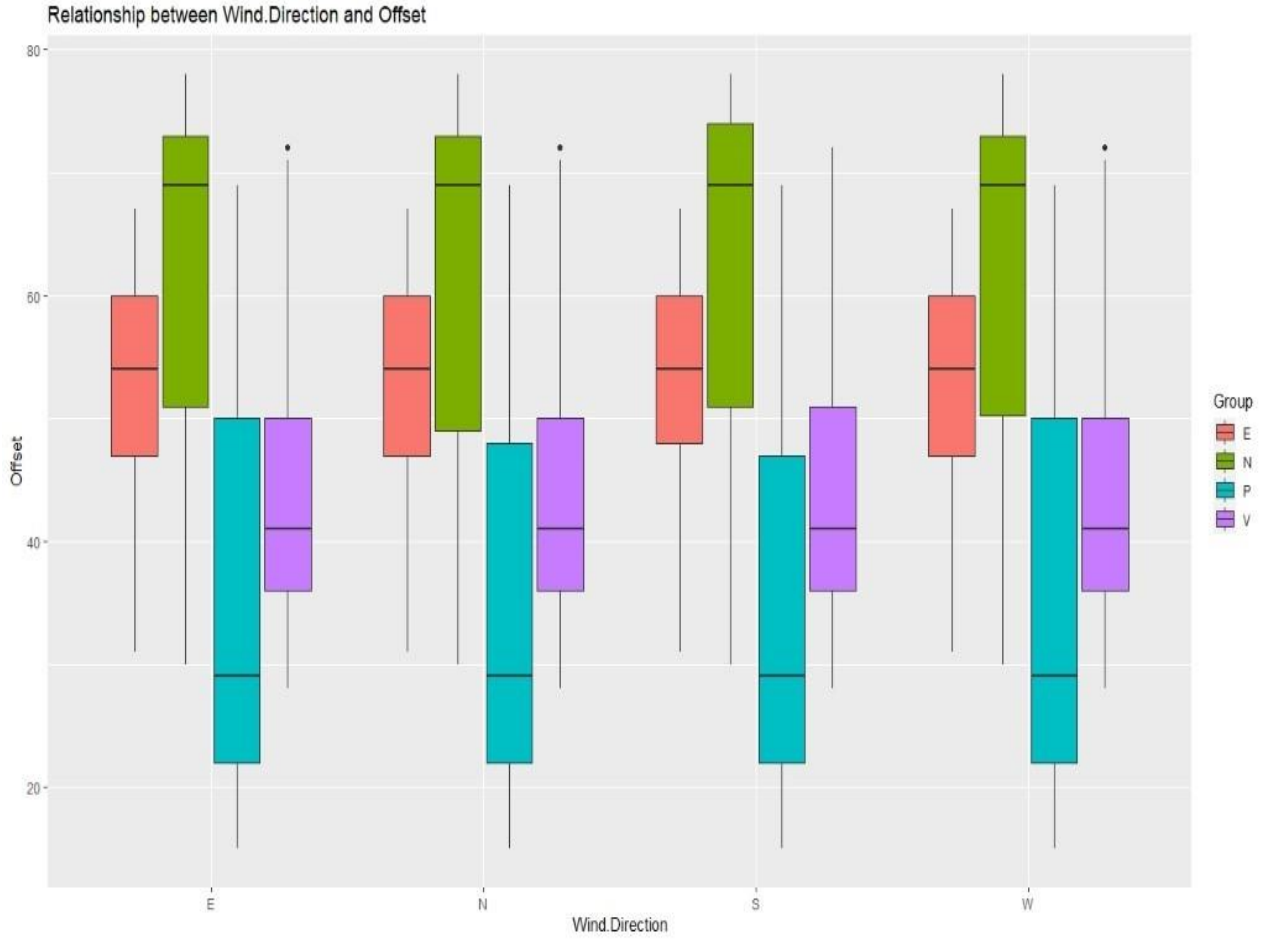

Fig. 6 Relationship between the Wind Direction and Offset.

;

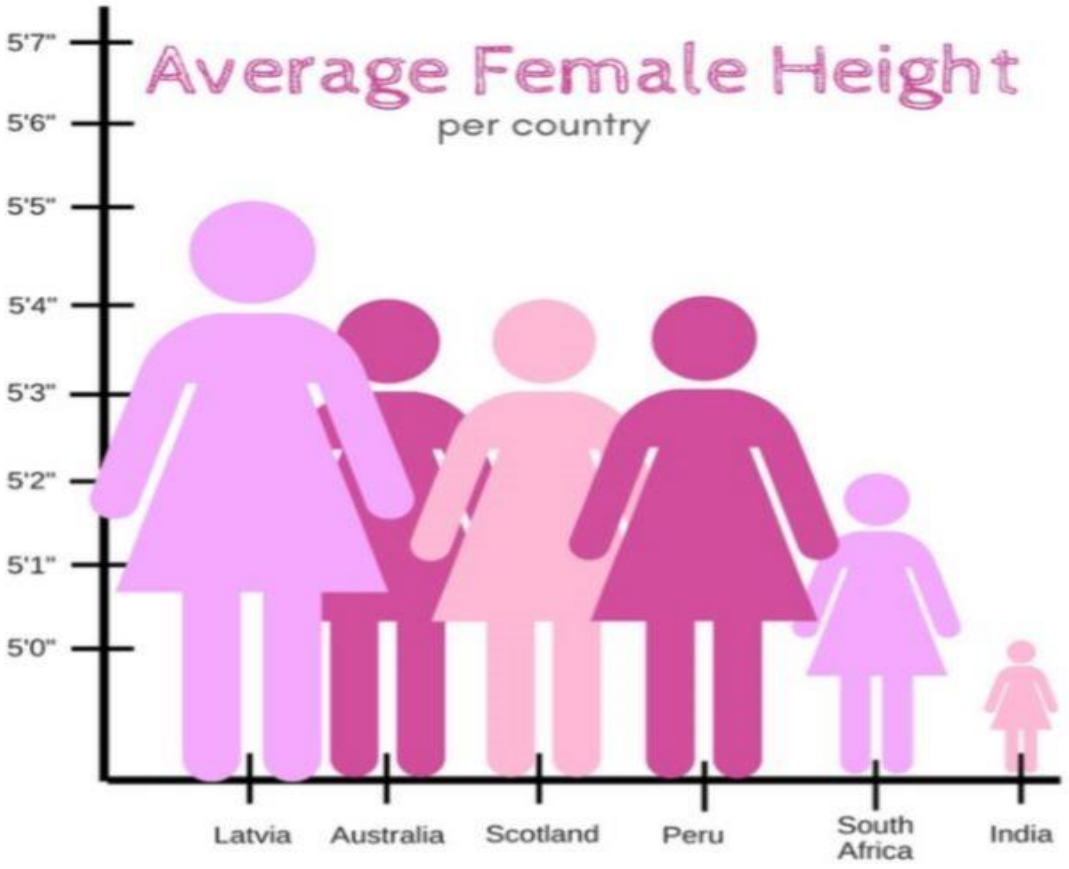

Fig. 8 Average Female Height Graph. Source: Coursework statement

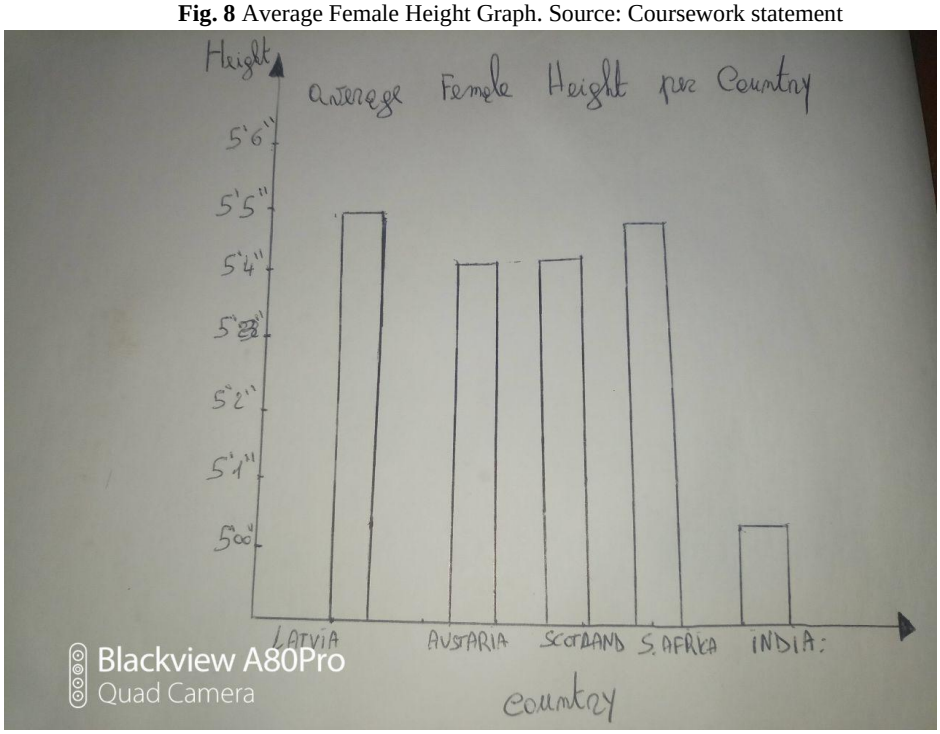

Fig. 9 Improved Average Female Height visualization. Realized using https://www.onlinecharttool.com/

4. References

1. Author, Davide Pollicino
2. Online free char generator: https://www.onlinecharttool.com/