Foundations of Statistics Class 1

Name of the Writer

Name of the University

Foundations of Statistics Class 1

**Introduction**

Use of statistical tools regarding the interpretation of a company's condition has been on the rise since forever. Moreover, the use of these tools can be implemented in all walks of life, which makes them an important branch of mathematics. Furthermore, its tools for interpretation, presentation, analysis, and preservation of the data regarding any business is pivotal in the management of a business. Statistical tools are important in rating organizational performance related to their finances, the market they are operating and their customers (Allegue, et al, 2017). The use of statistical tools is limitless in businesses and if they are implemented with proper skill, they can help in increasing the productivity within any corporation. Moreover, statistical tools can be very beneficial in terms of effective work management and increasing the performance of the employees. Furthermore, the health of any organization can be represented using these tools. In this context, this essay uses statistical calculations to tell the benefit of the data regarding the number of houses sold by Goodyear and how it can be beneficial for other companies as well.

**Importance of Statistical data**

This essay focuses on these specific statistical tools such as mean, median, standard deviation, and frequency distribution to find the benefit of the data regarding the number of houses sold by Goodyear estate agency.

**Mean**

|  |  |
| --- | --- |
| 229 | Mean |

The mean for the data is the measure of the central tendency of any data set and is applicable in both discrete and continuous data. The mean is the value that is the most common in the data and is not always one of the values that are listed in the data table. Note the sample mean of the data set of Goodyear real estate agency for the number of households sold is around two hundred and twenty-nine thousand dollars (Schabenberger and Gotway, 2017). This sale price reflects that at an average a household within that area is sold at this amount. Now, this value if important for agencies such as Goodyear and other companies is that when the customers who would want to buy property within that area, they will have an average selling price rate for them. They can compare this value with the budget of their customers and will help in deciding, if buying land or property within this area seems like a suitable investment for them. So, through the use of the mean, they can classify areas based on the average amount of a price of a good house sold within that area and align them with the needs of their different customers.

**Median**

|  |  |
| --- | --- |
| 216 | median |

 The median can be construed as another way of calculating the middle of any data set including numerical numbers. Where the mean is the average of the data set, the median is essentially the middle value between a numerical data set that equally separates the numerical values on both sides of the data set. So, this can be construed that point within the data set that has an equal number of values above and below its centre value. So, in essence, it makes the median the middle of the data set. Now the median can be important for Goodyear and other companies because of its unchanging nature and true representation of the dataset (Ali and Bhaskar, 2016). This is because with households sold at a higher price in the data, the mean of the data set would be shifted and be influenced onto the side with higher values or if there are smaller values than the mean, they would shift to the other. In this matter, as the median is the middle of the data set, it cannot be swayed to one side or the other. Meaning this would be a better value for classifying areas for future customers who want to buy houses within this area. With the median, there would be an equal number of house prices above it and below it.

**Standard Deviation**

|  |  |
| --- | --- |
| 47 | Standard deviation |

Essentially, using the standard deviation companies such as Goodyear and other companies, one can find the difference or the average distance between the numerical values in the data set and the sample mean calculated. With a low standard deviation, the values in the data set are seen to be close to the calculated mean of the data set. Whereas, with a high standard deviation, the values in the data set are spread out over a large range of values. So, by using the standard deviation Goodyear and other companies can determine if the data collected is valid. Furthermore, the confidence within this data set can be seen using the standard deviation (Modarres, 2016). This tells companies such as Goodyear and others can determine the level of error within the data by calculating the expected standard deviation in the results thinking that if the same sales are made over multiple years. Furthermore, if Goodyear wants to determine how precise and accurate the calculated mean of the data set is, they can use the standard deviation to determine this. Furthermore, the standard deviation is also helpful in determining the risk associated with the price fluctuation of the houses being sold in this area. This can be construed that it can help manage the risk of a portfolio containing different properties being sold or used.

**Frequency distribution**

|  |  |
| --- | --- |
|  Frequency Table  |   |
| Labels | Frequency |
| 100-150 | 1 |
| 150-200 | 15 |
| 200-250 | 20 |
| 250-300 | 10 |
| 300-350 | 5 |

Now the data that has been collected by Goodyear estate agency is for the prices that those houses were sold for and was very spread out. The data had no formal order for numerical values, which makes it harder for anyone to make sense of the data especially companies like other Goodyear estate agency. To formalize the data in a set pattern that could be understood by multiple parties the use of frequency distribution can occur. So, a frequency distribution table or just the frequency is used to summarize nominal ordinal data categorically (Sharp and Borkowski, 2016). Furthermore, it helps summarize continuous data when the entire data has been broken down into groups that are meant to be understood. The above table is the frequency table for the first fifty points of data collected by the Goodyear estate agency. Just by looking at the data, one can understand the number of households that were sold within any certain range. The highest number of the houses was sold in the price range of two hundred thousand to two hundred and fifty thousand dollars. This way the frequency distribution table shows where the data is clustered the most gives a better visual interpretation of the data while determining the different range of selling prices and finally focusing on the extreme values. In this case, the lowest extreme value is a hundred and forty-five thousand and the highest extreme value is three hundred forty-five thousand.

**Histogram**

 On the other hand, where the frequency distribution table is the numerical representation of the frequency of the data collected by the Goodyear estate agency, the histogram is the graphical representations of that data. The histogram takes the values for the ranges calculated in the above frequency distribution table and presents them in a bar chart form. This will be helpful for Goodyear estate agency and other companies as they can compare the different price ranges that the houses were sold at. The other companies could use this data to estimate the highest level of house pricing that can occur within this area.

**Conclusion**

 **I**t can be said that the use of statistical tools can have a huge impact on the daily workings of any business or an individual’s life. Through the use of statistical tools such as mean, median, standard deviation, frequency distribution, and a histogram, uncorrelated data can be summarized in such a way that it could be understood by an amateur user as well. Furthermore, each tool of statistics has its benefit of use as the mean tells us the average selling price. The median tells us the middle value of the data, the standard deviation locates the error in the data, frequency distribution classifies the data and the histogram is the graphical representation of this classified data.

**References**

Allegue, H., Araya‐Ajoy, Y. G., Dingemanse, N. J., Dochtermann, N. A., Garamszegi, L. Z., Nakagawa, S., & Westneat, D. F. (2017). Statistical Quantification of Individual Differences (SQuID): an educational and statistical tool for understanding multilevel phenotypic data in linear mixed models. *Methods in Ecology and Evolution*, *8*(2), 257-267.

Ali, Z., & Bhaskar, S. B. (2016). Basic statistical tools in research and data analysis. *Indian journal of anaesthesia*, *60*(9), 662.

Modarres, M. (2016). *Risk analysis in engineering: techniques, tools, and trends*. CRC press.

Schabenberger, O., & Gotway, C. A. (2017). *Statistical methods for spatial data analysis*. Chapman and Hall/CRC.

Sharp, J. L., & Borkowski, J. J. (2016). A conditional frequency distribution test for analyzing 2× c tables. *Statistics & Probability Letters*, *119*, 334-343.