Name Alden Butzke

|  |
| --- |
| 58 |
| 63 |
| 79 |
| 56 |
| 84 |
| 25 |
| 72 |
| 69 |
| 79 |
| 44 |
| 57 |
| 68 |
| 49 |
| 62 |
| 56 |

**BA 355: Business Analytics ICE 5**

1) Use Tukey’s method to determine the range of typical points and which points are outliers for the data set on the right.

1. What is the “typical” range?

**IQR = Q3 – Q1**

**14.5 = 70.5-56, 50% of the data falls between Q1 and Q3, and 14.5 is the range of the middle 50%, from 56 to 70.5. However, Tukey developed a method of finding outliers that goes beyond the highly data-sensitive nature of the above analysis using IQR.**

**Find Turkey’s method below.**

**[Q1 - 1.5\*IQR, Q3 + 1.5\*IQR]**

|  |  |
| --- | --- |
| **Lower** | **33.75** |
| **Upper** | **93.75** |

**The majority of the data should fall within this interval.**

1. Are there any outliers? If so, which data point(s) are they?

**Yes, 25 is an outlier because it falls beyond the Tukey test graphically shown by the box plot.**

2) Apply Tukey’s method to the 10,000 credit scores in column B of Case 2. Which range of credit scores are “typical” and which credit scores are outliers?

**The IQR (range of the middle 50% of the data) is 126, and the interval is from 644 (Q1) to 773 (Q3).**

**Any Credit score above 850 (max of data, though prediction is 950) or below 450 is an outlier by the Tukey test and box plot. There are 416 points less than or equal to 450.5.**

3) Apply Tukey’s method to the home prices for each data set you created for Case 3.1. Did you have any outlier homes in Durango, and were they on the low or the high end? How about in your other location?

**I had one outlier on the high end at $2.4 million. In Anchorage, I had an outlier at $1.79 million.**

**Anchorage:**

|  |  |
| --- | --- |
| Q1 | 382500 |
| Q3 | 650991.5 |
| IQR | 268491.5 |
|  |  |
| Lower | -20237.3 |
| Upper | 1053729 |
| **Durango:** |  |
| |  |  | | --- | --- | | Q1 | 608840.8 | | Q3 | 987928 | | IQR | 379087.3 | |  |  | | Lower | 40209.88 | | Upper | 1556559 | |  |