|  |
| --- |
| Describe Population Distribution |

**Population: (**Describe population with all known details**)**

**Variable:** **(**Give variable exactly as shown in source**)**  **Type:** Categorical or Quantitative, Measurement Scale, Discrete or Continuous

**Shape:** A (boxplot; dotplot; stemplot; and histogram) (were/was) examined to determine the shape of the distribution. The (dotplot; stemplot; and histogram) (were/was) displayed using a bin width of \_\_\_\_\_\_\_\_ (unit of measure) increments.

 All of these plots (were/was) found to be (**\_\_\_**modal) and (nearly symmetric; highly skew right; highly skew left; slightly skew right; slightly skew left).  **OR** The (boxplot; dotplot; stemplot; and histogram) appear(s) to be (**\_\_\_**modal) and (nearly symmetric; highly skew right; highly skew left; slightly skew right; slightly skew left). However, the (boxplot; dotplot; stemplot; and histogram) appear(s) to be (**\_\_\_**modal) and (nearly symmetric; highly skew right; highly skew left; slightly skew right; slightly skew left).

 The Fisher skew statistic was **\_\_\_\_**. This statistic fell (outside the computed range of **-\_\_\_\_\_\_ to +\_\_\_\_\_\_\_** **;** inside the computed range of **-\_\_\_\_\_\_ to +\_\_\_\_\_\_\_** **;** near 0) indicating that the distribution’s shape is (highly skew right; highly skew left; slightly skew right; slightly skew left; nearly symmetric).

**Center:** Mean = **\_\_\_\_\_\_** (units), Median = **\_\_\_\_\_\_** (units), Mode = **\_\_\_\_\_\_** (units)

 The best measure of central tendency is the (mean, median) because the distribution is (symmetric; skewed). This (symmetric; skew right; skew left) shape causes the mean to be (nearly equal to; greater than; less than) the median.

**Spread:** Range = **\_\_\_\_\_\_** (units), IQR = **\_\_\_\_\_\_** (units), **\_\_\_\_\_\_**(units)

 The best measure of spread is the (range and interquartile range; standard deviation) because the distribution is (symmetric; skewed).

**Outliers:** IQR Method: Adding 1.5 times the IQR to the third quartile value of **\_\_\_\_** (units) results in an upper outlier threshold of **\_\_\_\_\_** (units). Subtracting 1.5 times the IQR from the first quartile value of **\_\_\_\_\_** (units) results in a lower outlier threshold of **\_\_\_\_\_** (units). Examination of the data found (no; #) outliers that exceeded these thresholds, (list ALL outlier values).

 Standard Deviation () Method: Adding and subtracting three standard deviations from the mean of **\_\_\_\_\_** (units) establishes an upper outlier threshold of **\_\_\_\_\_\_** (units) and a lower threshold of **\_\_\_\_\_** (units). Examination of the data found (no; #) outliers that exceeded these thresholds, (list ALL outlier values).

 The best measure of outliers is the (IQR Method; Standard Deviation Method) because the distribution is (symmetric; skewed).

KEY: RED items should be chosen from options provided exactly as written. Note each option is separated by commas.

 BLUE items require appropriate words, usually the unit of measure (eg feet, $Millions, siblings)

 ORANGE lines should be filled with appropriate numbers.

|  |
| --- |
| Compare Two Distributions |

**Variable:** (Give variable exactly as shown in source) in the population of(Describe with all known details)

**Variable:** (Give variable exactly as shown in source) in the population of(Describe with all known details)

|  |  |  |  |
| --- | --- | --- | --- |
|  | **(Variable One Summary)** | **(Variable Two Summary)** | **Comparison** |
| **Shape** | (Near symmetric, slightly skew right, slightly skew left, highly skew right, highly skew left) | (Near symmetric, slightly skew right, slightly skew left, highly skew right, highly skew left) | The distributions have the same shape. **OR** The distribution for (name the distribution) is (nearly symmetric, highly skew right, highly skew left, slightly skew right, slightly skew left) while the distribution for (name the other distribution) is (nearly symmetric, highly skew right, highly skew left, slightly skew right, slightly skew left). |
| **Center** | Mean = **\_\_\_\_\_** (units)Median = **\_\_\_\_** (units) | Mean = **\_\_\_\_\_** (units)Median = **\_\_\_\_** (units) | Since both distributions are nearly symmetric, the best measure for comparing central tendencies is the mean. **OR** Since (both distributions are; the distribution for (name distribution) is) skewed, the best measure for comparing central tendencies is the median. The center of the distribution for (name the distribution) is (nearly the same; about \_\_\_\_ units higher; about \_\_\_\_\_ units lower) than the distribution for (name the other distribution).  |
| **Spread** | Range = \_\_\_\_\_\_IQR = \_\_\_\_\_\_\_\_\_\_\_\_ | Range = \_\_\_\_\_\_IQR = \_\_\_\_\_\_\_\_\_\_\_\_ | Since both distributions are nearly symmetric, the best measure for comparing spread is the standard deviation. **OR** Since (both distributions are; the distribution for (name distribution) is) skewed, the best measure for comparing spread are the range and interquartile range. Examination of these statistics shows both distributions have similar spreads. **OR** Examination of these statistics shows the distribution for (name the distribution) has more spread than the distribution for (name the other distribution).  |
| **Outliers** | (List outlier values) using the (IQR Method, Standard Deviation Method) | (List outlier values) using the (IQR Method, Standard Deviation Method) | Neither distribution has outliers. **OR**The distribution for (name the distribution) has **\_\_\_\_\_** outliers while the distribution for (name the other distribution) has none. **OR**The distribution for (name the distribution) has **\_\_\_\_\_** outliers, while the distribution for (name the other distribution) has **\_\_\_\_\_** outliers. |