

Business Statistics

Introduction

- Business Stats: provide the tool which such data are collected, analyzed, summarized, and presented to facilitate the decision making process
- Population: collection of persons or objects of interest “all”
- Census: data collected from the whole population
- Sample: a portion of the whole
- Parameter: a descriptive measure of the population
- Statistic: a descriptive measure of the sample
- Descriptive stats: using data gathered on a group to describe or reach conclusion about the same group
- Inferential stats: gathering data from a sample using the stats gathered to reach conclusions
- Variable: a characteristic of any entity being studied that is capable of taking on different values
- Ratio: have an absolute zero, same as interval
 - EX: height, weight, time, volume
- Interval: distances between consecutive numbers have meaning
 - EX: temperature, percentage change
- Ordinal: rank or order people and objects
 - EX: customer satisfaction survey
- Nominal: classify or categorized; no order
 - EX: religion, sex, geographic location, zip code
- Graphic picture: clear, concise, and consistent
- Ungrouped data: raw data that have not been summarized in any way
- Grouped data: data that have been organized into a frequency
- Frequency: a summary of data presented in the form of class intervals and frequencies
- **Class Width: range / # of classes**
- **Range: difference between largest and smallest numbers**
- **Midpoint: (high interval + low interval) / 2**
- **Relative frequency: individual class frequency / total frequency**
- **Cumulative frequency: the frequency for that class interval added to the preceding cumulative total**
- Histogram: a series of contiguous rectangles that represent the frequency of data
- Frequency polygon: a graphical display of class frequencies using the class midpoint
- Ogive: cumulative frequency polygon
- Dot plot: each data value is plotted along the horizontal axis and is represented by a dot

Business Statistics

- Pie chart: circular depiction of data where the area of the whole pie represents 100% of the data and slices of the pie represent a percentage change breakdown of the sub levels
- Scatter Plot: two dimensional graph plot of pairs of points from two numerical variables

Business Statistics

Descriptive Statistics

- Mode: most frequently occurring value in a set of data
- Mean: average
- Median: middle value / average of the middle two terms
- Percentiles: measures of central tendency that divide a group of data into 100 parts
 - $I = P/100 (N)$
- Quartiles: measures of central tendency that divide a group of data into four segments
 - $Q1 = P25 = (\text{location of } I + \text{right of it}) / 2$
- Range: difference between the largest value of a data set and the smallest value of a set
- Interquartile range: range of values between the 1st and 3rd quartile
- Mean absolute deviation: the average of the absolute values of the deviations around the mean of a set of numbers
 - $MAD = (\text{summation of } xi - \text{meau}) / n$
- Variance: the average of the squared deviations about the arithmetic mean for a set of numbers
 - $(\text{summation of } xi - \text{meau})^2 / N$
- Standard deviation: the square root of the variance
 - $\text{Square root}[\text{ summation of } xi - \text{meau})^2 / N]$
- Z-score: represents the number of standard deviations a value is above or below the mean of the mean of a set of numbers when the data are normally distributed
 - $Z = (xi - \text{mean}) / \text{standard deviation}$
- Coefficient of variation: the ratio of the standard deviation to the mean expressed in percentage
 - $CV = (\text{standard deviation} / \text{meau}) \times 100$

Business Statistics

Probability

- Probability: the ratio of the number of items in a population containing the event to the total number of items in the population
 - $P(E) = \text{total number of outcomes} / \text{total number of possible outcomes}$
 - $P(E) = \text{number of times an event occurred} / \text{total number of opportunities}$
- Experiment: a process that produces outcomes
- Event: an outcome of the experiment
- Elementary event: an event that cannot be decomposed or broken down into other events
- Sample space: a complete roster of listing of all elementary events for an experiment
- Union(U): union of x and y contains the elements from each of the sets, **Combine all, no duplicates**
- Intersection(Backwards U): intersection of x and y contains the elements common to both sets, **DUPLICATES**
- Mutually exclusive events: if the occurrence of one event precludes the occurrence of the other event
- Independent events: if the occurrence or nonoccurrence of one of the events does not affect the occurrence or nonoccurrence of the other event
- Complementary events: all elementary events of an experiment not in x comprise its complement
- Marginal probability: computed by dividing some subtotal by a whole
- Union probability: probability that E1 will occur or E2 will occur
- Joint probability: the probability that E1 and E2 are occurring
- Conditional probability: probability that E1 will occur given that E2 is known to have occurred
- General law of addition: used to find the probability of the union of two events
 - $P(X \text{ union } Y) = P(X) + P(Y) - P(X) \times P(Y)$
- Special law of addition: if two events are mutually exclusive, the probability of the union of the two events
 - $P(X \text{ union } Y) = P(X) + P(Y)$
- General law of multiplication: gives the probability that both events X and Y will occur at the same time
 - $P(X \text{ backwards } U Y) =$
 - $P(X) \times P(Y|X)$
 - $P(Y) \times P(X|Y)$
- Conditional probability: probability that X will occur given Y
 - $P(X|Y) = P(X \text{ backwards } U Y) / P(Y)$

Business Statistics

Distributions

- Random variable: a variable that contains the outcomes of a chance experiment
 - Discrete if countably infinite
 - Continuous if values are at every point over a given interval
 - EX: time, height, weight, and volume
- Mean: summation [$X \cdot P(X)$]
- Variance: summation [$(X - \text{meau})^2 \cdot P(X)$]
- Binomial distribution:
 - N: number of trials
 - X: number of successes
 - P: probability of success
 - Q: probably of failure
 - $P(X) = \binom{n}{x} p^x q^{n-x}$
 - $\text{Meau} = np$
 - $\text{Variance} = npq$
 - $\text{Standard deviation} = \text{square root}(npq)$
- Poisson distribution: $P(X) = \frac{\text{Average}^x e^{-\text{Average}}}{x!}$
 - Mean: average
 - Variance: average
 - $\text{Standard deviation} = \text{square root}(\text{average})$
- Uniform distribution: a continuous distribution in which the same height is obtained over a range of values
 - $\text{Height} = 1 / b - a$
 - $\text{Length} = b - a$
 - $\text{Meau} = a + b / 2$
 - $\text{Standard deviation} = b - a / \text{square root}(12)$
 - $\text{Probability} = P(X) = X2 - X1 / b - a$
- Normal distribution: continuous, symmetrical about its mean, unimodal, area = 1
- $Z = x - \text{meau} / \text{standard deviation}$

Business Statistics

Problem Set 1 Notes

- At least \$100: Above \$100
- Under \$100: \$100 and under
- Variance: $(\text{Value} - \text{mean})^2 / n - 1$
- Standard deviation: Square root of Variance
- $P(X | Y)$: $P(X \text{ backwards } U Y) / P(X)$
- $P(Y | X)$: $P(X \text{ backwards } U Y) / P(Y)$
- Either or: $P[M \cup (x < 20)]$: $P(M) + P(x > 20) - P[M \text{ backwards } U (x > 20)]$
- Mutually Exclusive: $P[(x < 20) \cup (x > 25)]$: $P(x < 20) + P(x > 25) - 0$