github.com/HACN85

DATA ANALYSIS MADE SIMPLE:

EMBRACING STATISTICAL CONCEPTS WITH EXAMPLES

Swipe →

github.com/HACN85

PROBABILITY + BAYES' THEOREM

Probability assesses the likelihood of events. Bayes' Theorem updates our belief about an event based on prior knowledge. Think of predicting the probability of rain tomorrow based on historical data and current weather patterns.

Swipe →

github.com/HACN85

DISCRETE & CONTINUOUS RANDOM VARIABLES

Discrete variables (like rolling a dice) have specific, separate values. Continuous variables (like time or weight) can take any value within a range. For instance, in Python, generating random numbers using random.randint() (discrete) and random.uniform() (continuous) showcases this.

Swipe →

github.com/HACN85

SAMPLING DISTRIBUTIONS

When we collect multiple samples and calculate a statistic (mean, standard deviation), these values form a distribution. In Excel, using NORM.INV() to simulate a sampling distribution demonstrates this.

Swipe →

github.com/HACN85

PARAMETERS & ESTIMATORS

Imagine you're a teacher and want to know the average score of all students in your school, but you can't ask every student. So, you ask 30 random students for their scores and find their average score. That average score from the 30 students is your best guess (an estimator) for the actual average score of all students in the school (a parameter). The parameter is what you want to find, and the estimator is your calculated guess based on the smaller group you surveyed. This concept of using a sample to estimate characteristics of a larger group is key in statistics.

Swipe →

github.com/HACN85

CENTRAL LIMIT THEOREM

It states that the distribution of sample means approaches a normal distribution, regardless of the original distribution. Imagine flipping a coin multiple times and observing the distribution of heads - it tends towards a bell curve.

Swipe →

github.com/HACN85

CONFIDENCE INTERVALS

They provide a range of values we're confident the true parameter lies within. Calculating a confidence interval for the mean using Excel's CONFIDENCE() function exemplifies this.

Swipe →

github.com/HACN85

HYPOTHESIS TESTING

It's about making inferences based on data and testing assumptions about a population parameter. Conducting a t-test in Python (e.g., scipy.stats.ttest_ind()) illustrates this practice.

Swipe →

github.com/HACN85

TESTS FOR VARIANCE

Different statistical tests help us assess variations within data sets. For instance, using Excel's VAR.P() and VAR.S() functions to test variance within a dataset showcases this.

Swipe →

github.com/HACN85

WAS THIS HELPFUL?

DONT FORGET TO SAVE THIS POST