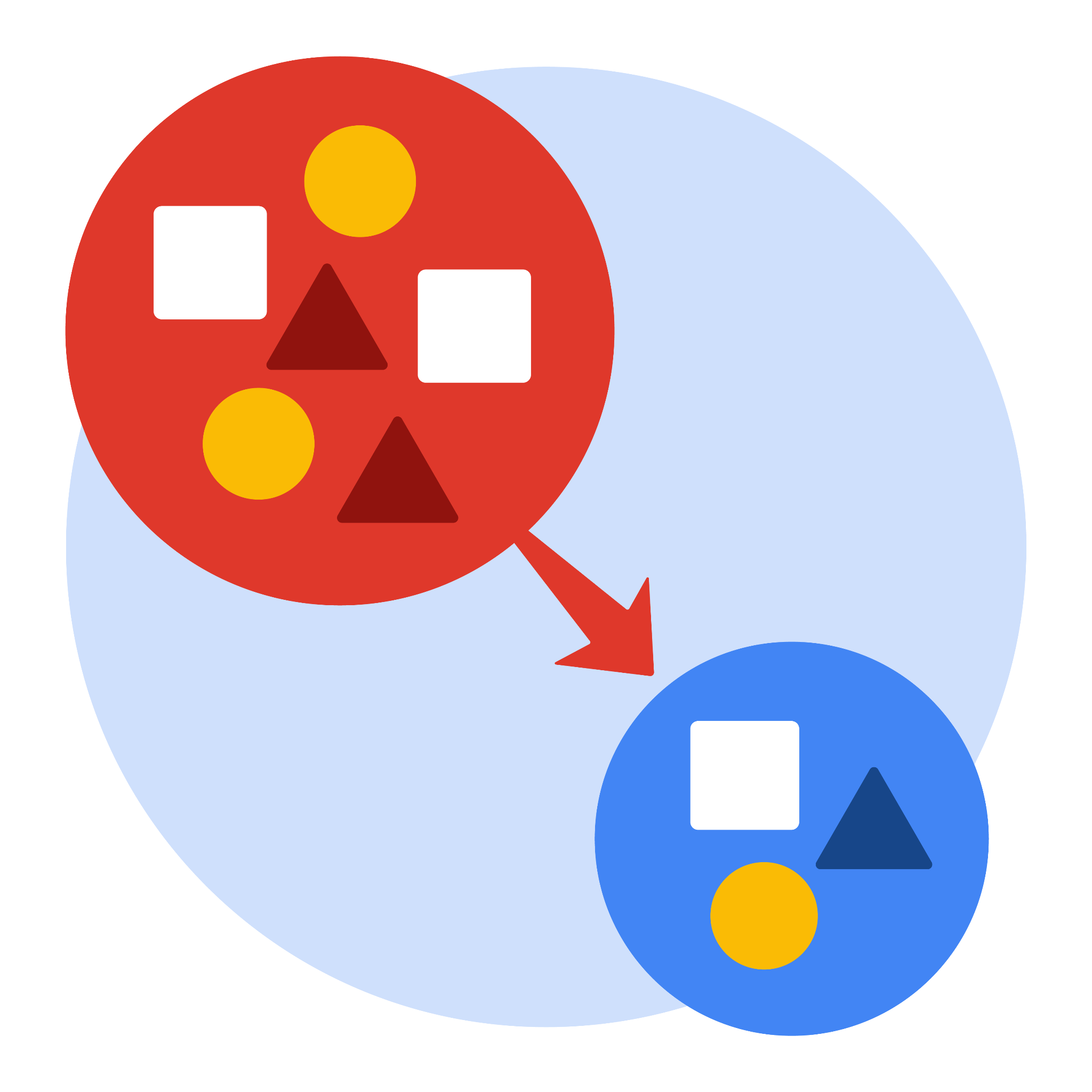
**Course Four**

# From Data to Insight: The Power of Statistics



# 

**Data Project Questions & Considerations**

**PACE: Plan Stage**

* What is the main purpose of this project?

To determine whether there is a statistically significant difference in taxi fare amounts between customers who pay by credit card and those who pay with cash, in order to inform TLC’s decision-making.

* What is your research question for this project?

Do customers who pay by credit card pay higher fares on average than those who pay with cash?

* What is the importance of random sampling?

Random sampling ensures that the groups being compared are representative of the overall population, which reduces bias and increases the validity of the results.

* Give an example of sampling bias that might occur if you didn’t use random sampling.

If credit card payments were only sampled during peak hours and cash payments during non-peak hours, fare differences might reflect time-of-day pricing rather than actual payment method effects.



 **PACE: Analyze & Construct Stages**

* In general, why are descriptive statistics useful?

They help summarize and understand the main features of a dataset, such as central tendency, variability, and distribution patterns.

* How did computing descriptive statistics help you analyze your data?

It allowed me to compare the average fare amounts and standard deviations for credit card vs. cash payments, revealing initial insights into fare differences.

* In hypothesis testing, what is the difference between the null hypothesis and the alternative hypothesis?
* The null hypothesis assumes no difference between the groups.
* The alternative hypothesis assumes there is a meaningful difference.
* How did you formulate your null hypothesis and alternative hypothesis?
* H₀: There is no difference in fare amount between credit card and cash users
* H₁: The fare amount is higher for credit card users than for cash users
* What conclusion can be drawn from the hypothesis test?

Since the p-value was effectively 0, we reject the null hypothesis. There is a statistically significant difference — credit card users pay higher fares on average than cash users.

**PACE: Execute Stage**

* What key business or organizational insight(s) emerged from your A/B test?

Customers who pay with a credit card tend to pay more on average than those who pay with cash, which may suggest a correlation with longer trips, tipping behaviors, or ride contexts.

* What recommendations do you propose based on your results?

TLC or taxi operators could implement strategies to promote credit card usage — such as rewards programs, discounts for using cards, or marketing efforts — to potentially increase fare revenues.