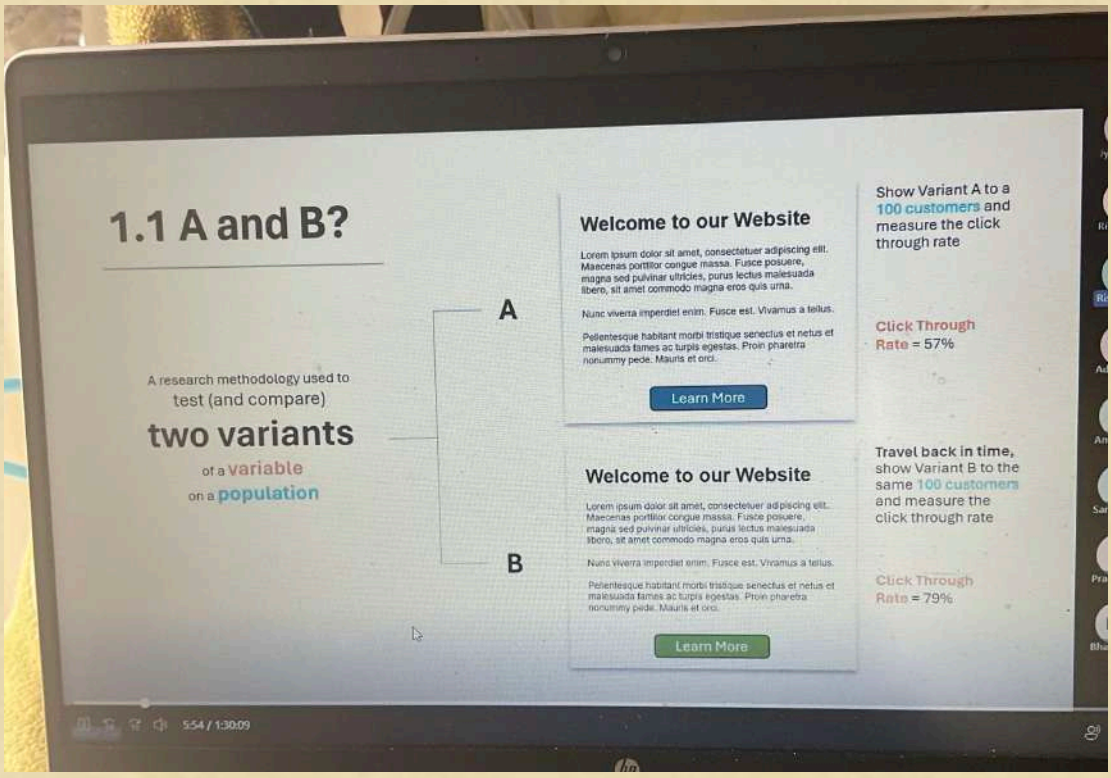
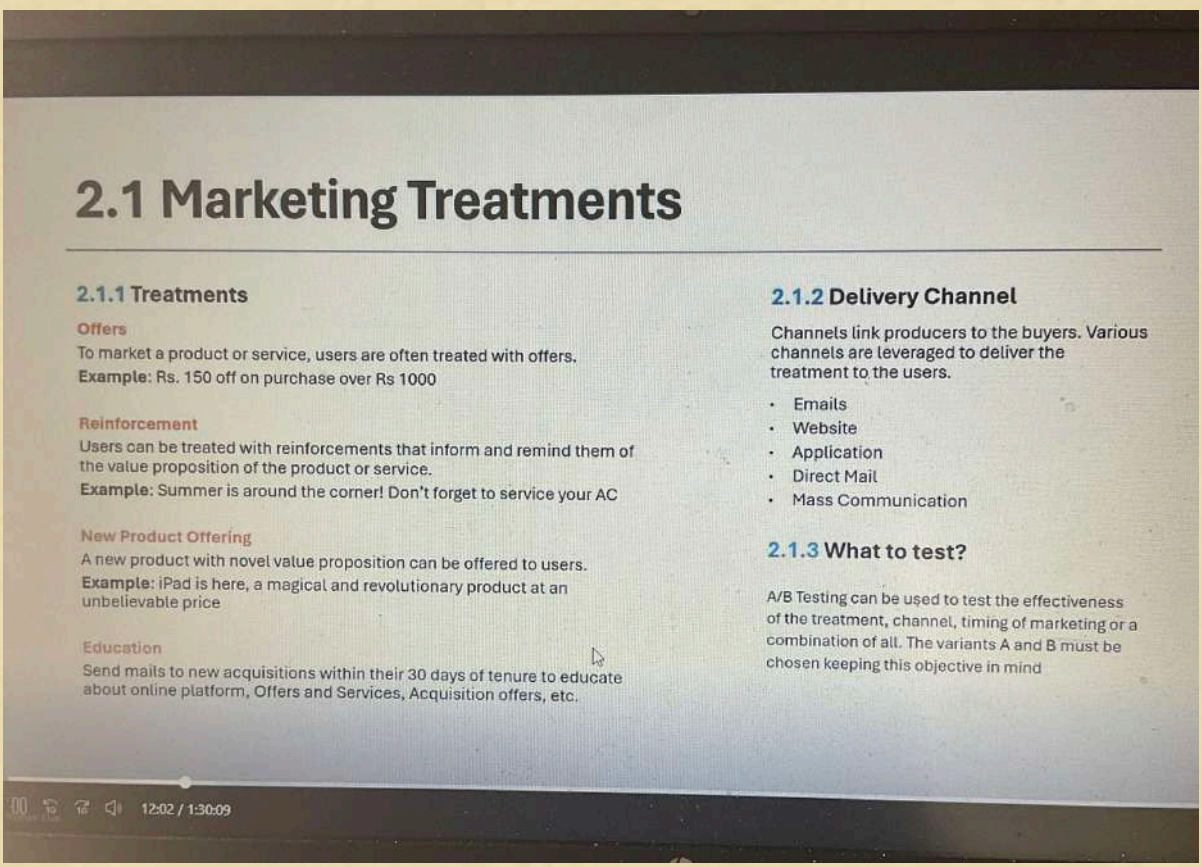
Ex L - 30 April


1.1 A and B?
A research methodology used to test (and compare)
two variants
of a variable
on a population
A
B
Welcome to our Website
Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas porttitor congue massa. Fusce posuere, magna sed pulvinar ultricies, purus lectus malesuada libero, sit amet commodo magna eros quis urna.
Nunc viverra imperdiet enim. Fusce est. Vivamus a tellus.
Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Proin pharetra nonummy pede. Mauris et orci.
Learn More
Show Variant A to a 100 customers and measure the click through rate
Click Through Rate = 57%
Welcome to our Website
Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas porttitor congue massa. Fusce posuere, magna sed pulvinar ultricies, purus lectus malesuada libero, sit amet commodo magna eros quis urna.
Nunc viverra imperdiet enim. Fusce est. Vivamus a tellus.
Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Proin pharetra nonummy pede. Mauris et orci.
Learn More
Travel back in time, show Variant B to the same 100 customers and measure the click through rate
Click Through Rate = 79%
5:54 / 1:30:09


2.1 Marketing Treatments
2.1.1 Treatments
Offers
To market a product or service, users are often treated with offers.
Example: Rs. 150 off on purchase over Rs 1000
Reinforcement
Users can be treated with reinforcements that inform and remind them of the value proposition of the product or service.
Example: Summer is around the corner! Don't forget to service your AC
New Product Offering
A new product with novel value proposition can be offered to users.
Example: iPad is here, a magical and revolutionary product at an unbelievable price
Education
Send mails to new acquisitions within their 30 days of tenure to educate about online platform, Offers and Services, Acquisition offers, etc.
2.1.2 Delivery Channel
Channels link producers to the buyers. Various channels are leveraged to deliver the treatment to the users.
Emails
Website
Application
Direct Mail
Mass Communication
2.1.3 What to test?
A/B Testing can be used to test the effectiveness of the treatment, channel, timing of marketing or a combination of all. The variants A and B must be chosen keeping this objective in mind
12:02 / 1:30:09

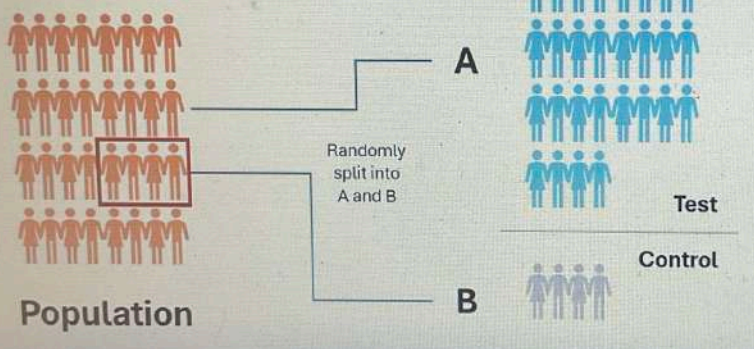
2.2 Test Control Framework

- Marketing team reaches out to their customers through campaigns and eNL.
- The eligible population is identified based on multiple parameters
- This eligible population is then divided into Test and Control groups (90/10, 95/5)

  Test- Population who received the mail

  Control- Population who didn't receive the email
- Both test and control group have similar behavior and should be mutually exclusive
- Test group could further be split based on creatives

2.3 Setup for Testing



2.2.1 Size of Control

The power of a hypothesis test is the probability that the test correctly rejects the null hypothesis when a specific alternative hypothesis is true.

$$n = f(s, d, \alpha, \beta)$$

where s is the population standard deviation, d is the target difference to detect, α is the required significance level and β is the power.

A large control group is usually not feasible as it restricts marketing reach

2.2.2 Ensuring Randomness

The division of test and control must be random – both groups should be representative of the population.

This can be a challenge in analysis involving smaller populations

$\alpha = 0.05$

$\beta = 0.80$

In real life companies want to spend more.

→ Power analysis is used to find min control sample so that test is statistically significant.

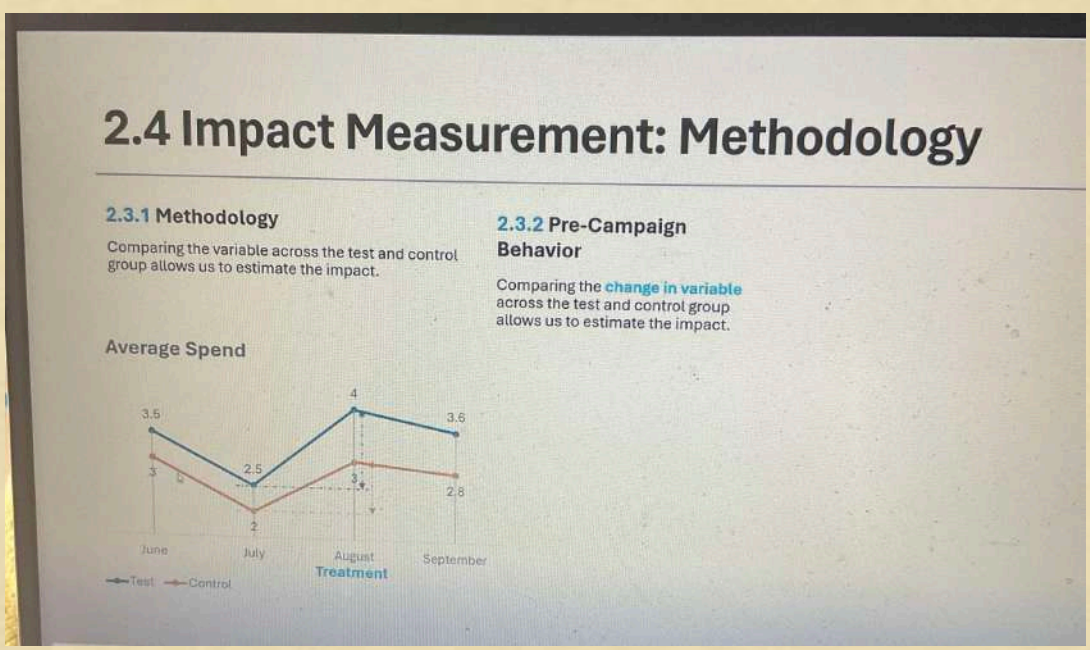
→ Reach as many people.
so large test group
so run campaign for longer

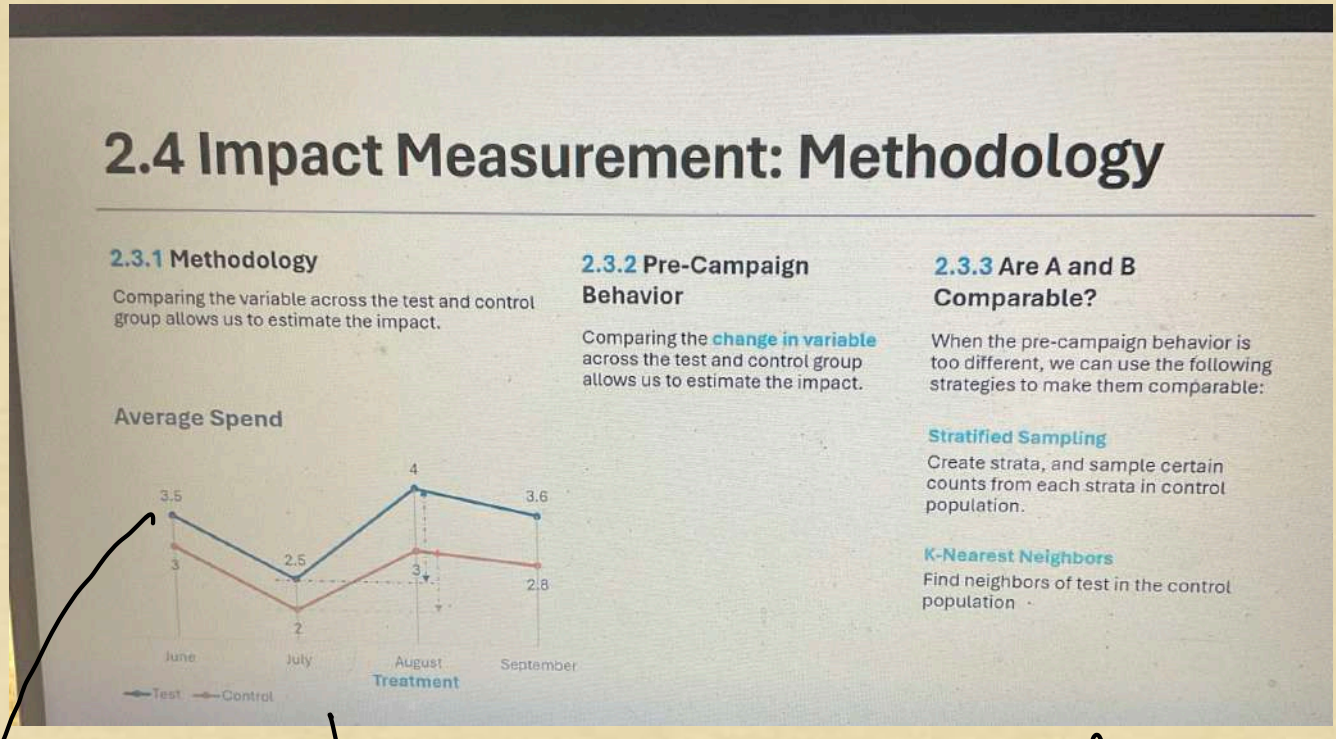
Market want → Smaller control group.

Statistician want → Larger control group for evidence

Test → 90 // control → 10 //

Both control & test — should be random.





↓

in control with high avg spend remove the highest quartile bucket. //

↓

so stratified sampling.

sampling → Only for control group

↓

stratified sampling or KNN is only for control group.

↓

this is done so control & test group are comparable

→ basically avg spend of control & test have diff of .5 //

2.5 Impact Measurement: Methodology

Methodology

Creating mutually exclusive Test & Control Population by removing overlapping cases.

1. Checking Pre email profile of Test & Control on multiple parameters -
2. If profiles are not similar, we create sampled control using stratified sampling.

Calculate KPI- For post and pre period from Mail Drop Month/Mail Month

Test Growth = Post Test Metrics – Pre Test Metrics
Control Growth = Post Control Metrics – Pre Control Metrics

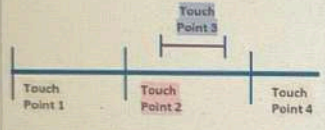
Impact per CM = Test Growth - Control Growth

Significance Testing

2.6 Campaign Setup: Deep Dive

Campaign Objective, Possible Roadblocks and Solutions

Customer Journey



Email
Digital Journey
Direct Mail

These campaigns are executed on a weekly/hourly/monthly basis

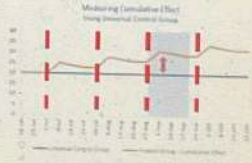
Campaign Objectives

Measuring Impact of Each Touch-Point

Touch point specific Holdout groups are required. These groups receive all other communications.

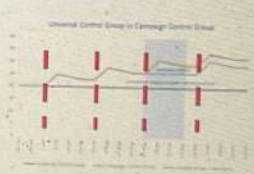
Measuring Holistic Impact of Campaign

A **Universal Holdout** that doesn't receive communication at any of the touch points is required.



Measuring Holistic Impact & Impact of Each Touch Point

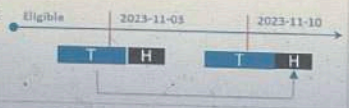
Both types of holdout groups will be required.



Possible Roadblocks

1. Holdout in 1 touchpoint (spills over to)/(is from) mailable of (subsequent)/(previous) touchpoints
2. Lack of distinction between universal and touchpoint specific holdout groups
3. Insufficient holdout groups exposes population averages to outlier effects.

Sample Journey



Possible Solutions

- **Manual Intervention-** Share retargeting instructions in the plan
- **Leverage Unique ID-** Identify variables whose any 1 or 2-digit numbers are randomly distributed for population. Pick different set of those numbers for mailable and holdout groups

↳ finding out holding out group

↓

understand Marketers — what is impact of mail on 10th day

Mails on 10th, 20th, 30th, 40th days

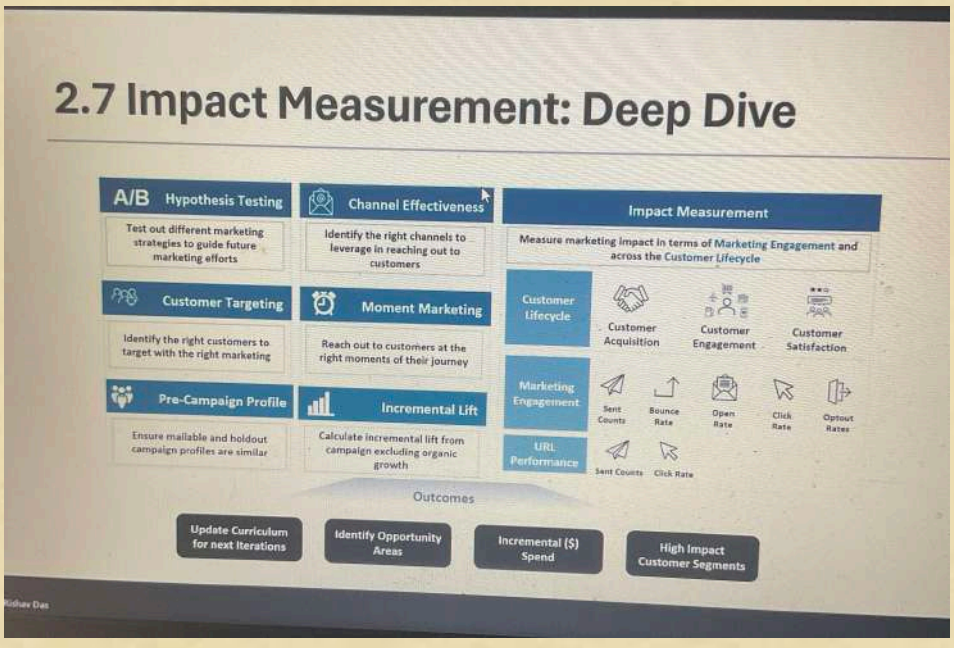
10th day
hold out group

— what is impact of all mails of 10th, 20th, 30th, 40th day?

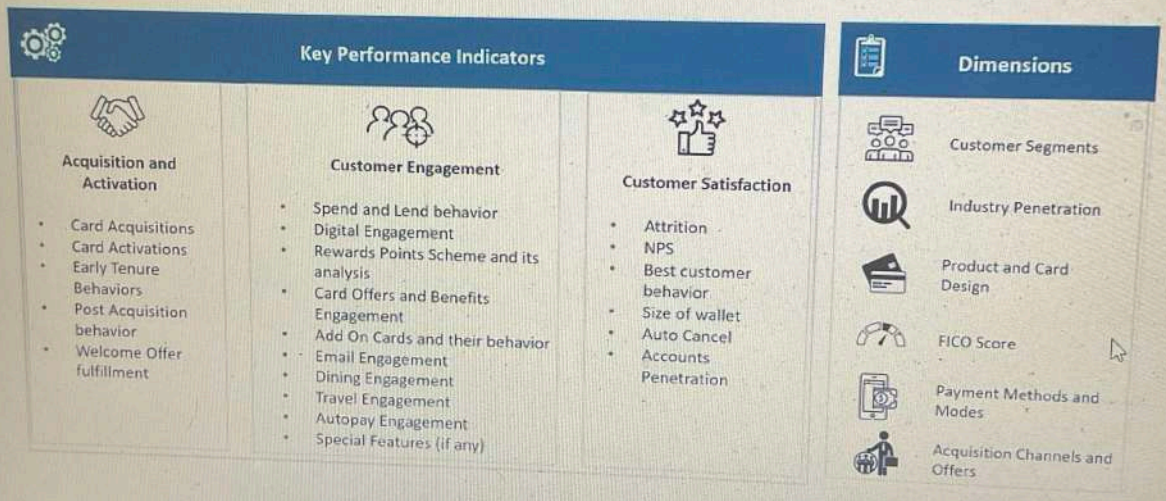
people who did not get any of the mail universal.

for population set up:
each user from no. from 1 – 99 →
all people ending with no. 5 → hold out group 1.
another set 2 → hold out group 2
another se 3 →

↳ basically control group



2.8 Credit Card Industry: KPIs



Exercise

1. Test and control population have their pre campaign average spend within 30%, pre campaign average Tenure within 20% and pre campaign average Credit Score is within 5%. Can we use test and control population as is?
   - A. Yes
   - B. No, we need to take additional steps to make test and control comparable
2. There is an overlap of some customers in test and control group, how should we handle the overlapping customers?
   - A. Include only in test
   - B. Include in both test and control
   - C. Include only in control
3. In one of the campaigns, customers received 1 mail per quarter. Same set of customers receive mail in Q1, Q2, Q3 and Q4. After end of campaign, we want to measure the overall impact of campaign. What types of hold out groups are required?
   - A. Touchpoint (Quarter) specific holdout
   - B. Universal Holdout
   - C. Both A and B

Campaign Execution Framework

To ensure a successful campaign execution, best practices must be followed at all individual stages

Campaign Design	❑ Set clear campaign objectives e.g. – Drive Activation, Influence Spend etc. ❑ Identify right channels ❑ Develop content in line with objectives
Population Selection (Targeting)	❑ Select target segments for the campaign ❑ Filter out population through Base suppressions/ Offer specific suppressions ❑ Prioritize target population based on Cost constraints /ROI target
Test Design	❑ Based on the campaign objectives, determine the metrics that will be measured (can be separated into **direct** and **downstream** metrics) ❑ Determine size of Test vs Control based on population constraints, nature of campaign, expected lift in key metrics and allowable significance level
List Pull and Validation	❑ Perform cross validation of population counts after list pull by comparing to counts used for population selection ❑ Ensure similarity of Test and Control by comparing certain key metrics like Overall spend, Peer Spend, Tenure, Other Demographics, etc.
Rollout	❑ Record any deviations from the campaign plan during the rollout process e.g. non-mailing to certain ZIP codes, technical issues with web impressions, etc. ❑ Record any special changes that may affect campaign results e.g. macroeconomic changes, incentive changes etc.
Analysis	❑ Compare relevant metrics across Test and Control to identify campaign impact ❑ If comparing other metrics than previously intended, ensure that the population is sufficient for obtaining a reliable measurement

Chapter 3: Statistical Significance of Impact: Hypothesis Testing

3.1 Introduction

3.1.1 A Coin Toss Experiment

We want to prove that a coin is biased towards heads. To do this, we toss the coin 10 times.

Coin 1

We get **6 heads and 4 tails**. Is this evidence strong enough to say that the coin is biased towards head?

$P(6, 10) = 0.377$

Coin 2

We get **9 heads and 1 tails**. Is this evidence strong enough to say that the coin is biased towards head?

$P(9, 10) = 0.011$

1. Assume that the coin is fair
2. What is the probability that we get at least n heads by chance?
3. Is that probability low enough for us to say that the coin is unfair?

Probability of Favorable Outcome

$P(n, 10)$ We calculate the probability of getting at least n heads in a **fair coin**, tossed ten times

C

3.1.2 The Hypothesis

We hypothesize that $\bar{X}_{test} > \bar{X}_{control}$

Observing a +ve impact is not sufficient evidence to support the hypothesis.

Population that received Treatment

90% Sample

Person	Spend
A	X_A
B	X_B
C	X_C
...	...
Average	X_{test}

Population that did not receive Treatment

10% Sample

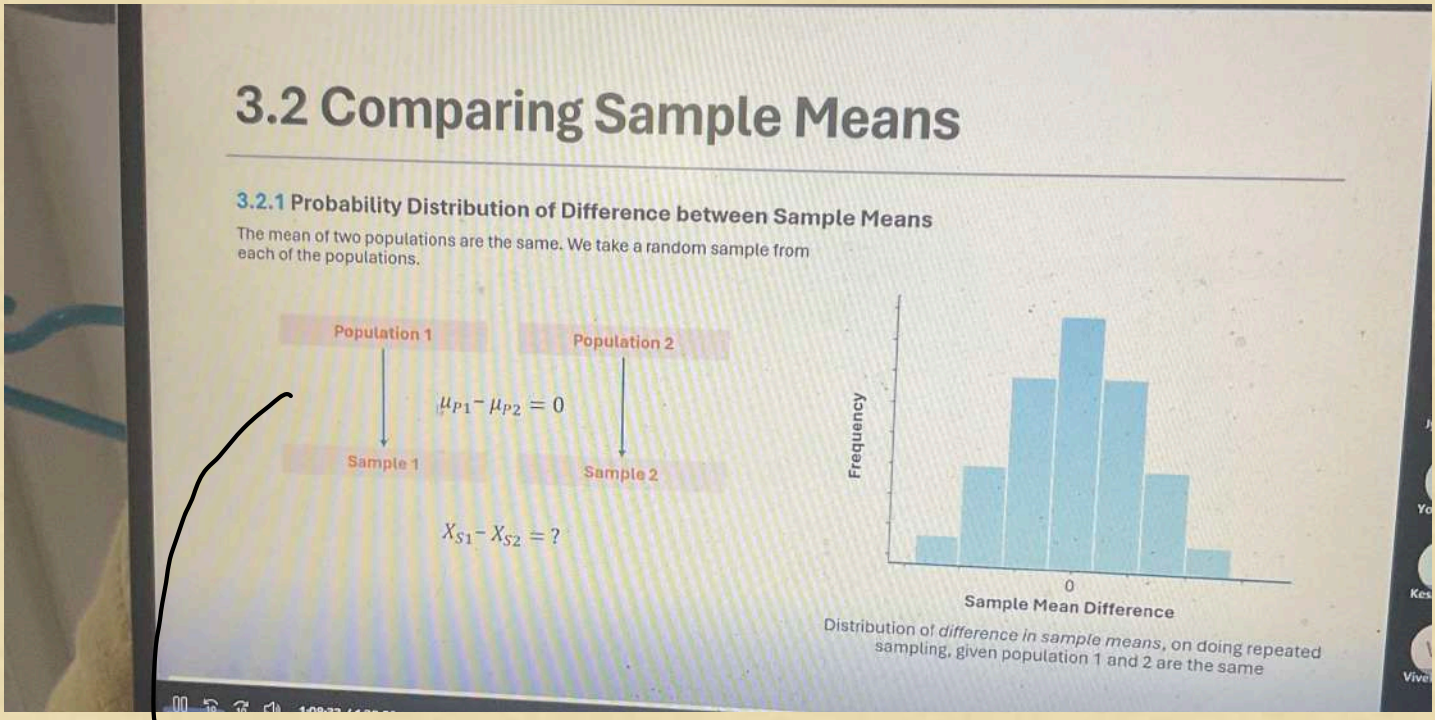
Person	Spend
P	X_P
Q	X_Q
R	X_R
...	...
Average	$X_{control}$

Impact of Campaign

$$Y = \bar{X}_{test} - \bar{X}_{control}$$

1. Assume that the population means are the same (H_0)
2. **What is the probability that we get at least Y impact by chance?** $P(Y \text{ given } H_0)$
3. Is that probability low enough for us to say that test mean is higher than control mean?

$$H_0: \quad \bar{X}_{test} = \bar{X}_{control}$$
$$H_A: \quad \bar{X}_{test} > \bar{X}_{control}$$



Actually

Population

|

Sample

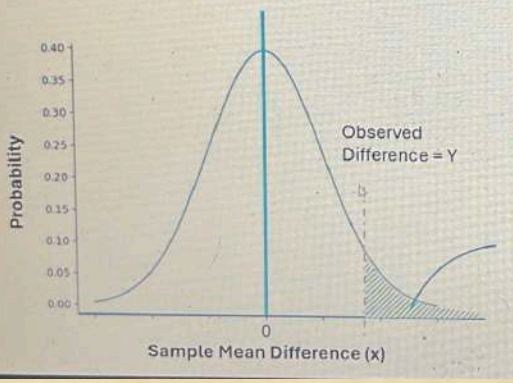
90% of Sample → Test

10% of Sample. — control

3.2.2 Probability of Observed Impact

Assume that the means of the two populations are the same.

The probability of observing a difference of x between the means of samples from these two populations follows a t-distribution.



P – value

Probability of observing a difference of Y or greater, if the populations means are same

If the probability of observing difference greater than Y is low (shaded in green), we can say that our assumption was wrong, implying enough evidence that the mean of the test population is greater than control population

3.2.3 How low is Statistically Significant?

Values of p	Inference
$p > 0.100$	No evidence against the null hypothesis
$0.050 < p < 0.100$	Weak evidence against the null hypothesis
$0.010 < p < 0.050$	Moderate evidence against the null hypothesis
$0.005 < p < 0.010$	Good evidence against the null hypothesis
$0.001 < p < 0.005$	Strong evidence against the null hypothesis
$p < 0.001$	Very strong evidence against the null hypothesis

Assumptions

- Data collected is continuous or ordinal
- The samples are randomly selected
- The population data follows normal distribution

3.3 Choosing the Test

Depending on the situation, appropriate test must be chosen.

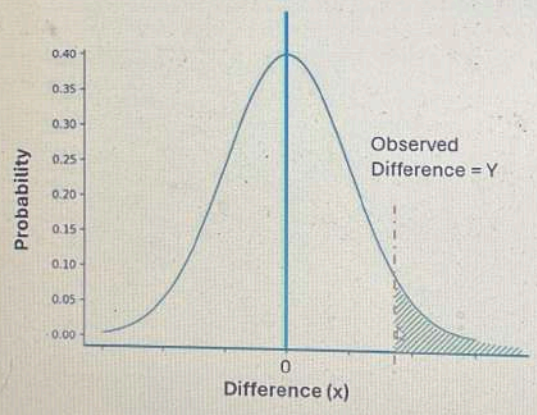
so far if working

Situation	Test
Difference between an assumed population mean μ0 and a sample mean $\bar{x}$,	
When the population variance is known	Z-test
When the population variance is unknown	t-test
Difference between two sample means, one from each population,	
When the population variances are known and equal	Z-test
When the population variances are known and unequal	Z-test
When the population variances are unknown but equal	t-test
When the population variances are unknown and unequal	t-test
When observations for the two sample are obtained in pairs	t-test
Difference between an assumed population proportion and an observed sample proportion	Z-test
Difference between two sample proportions, one from each population	Z-test
Difference between two counts	Z-test
Difference between an assumed population variance and a sample variance	χ^2-test
Difference between two sample variances, one from each population	F-test
A variable in Regression Model (i.e. difference between a regression coefficient and zero)	t-test
Overall regression model	F-test

Assumption = bell curve

3.4 One or Two Tailed?

3.3.1 One Tailed Test

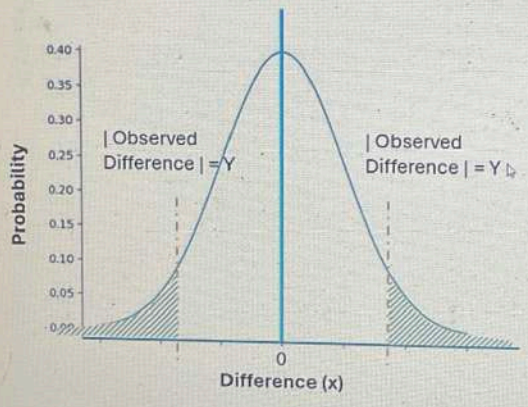


H_A: $\bar{X}_{test} > \bar{X}_{control}$

mean of test > mean of control

3.5 One or Two Tailed?

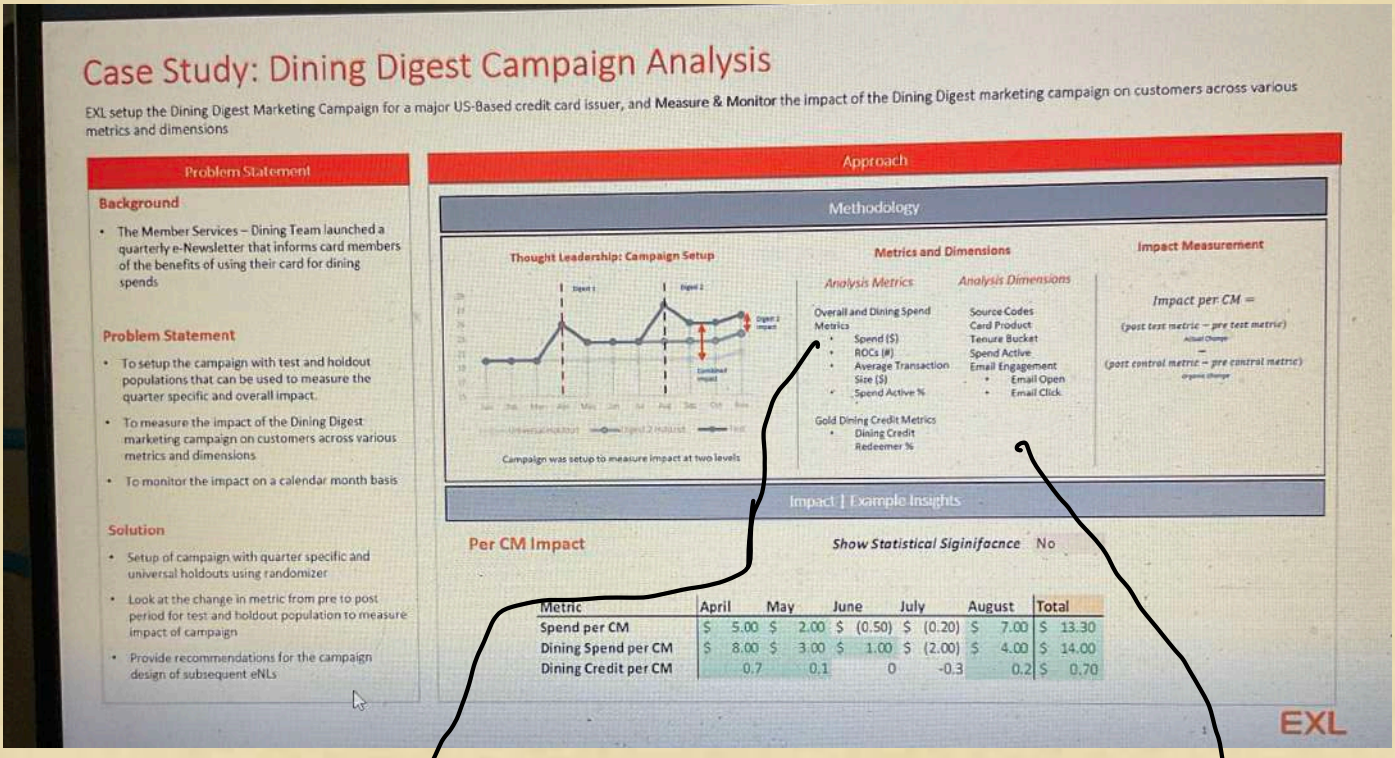
3.3.2 Two Tailed Test



H_A: $\bar{X}_{test} <> \bar{X}_{control}$

mean of test different from mean of control

Example.



Case Study: Dining Digest Campaign Analysis

EXL setup the Dining Digest Marketing Campaign for a major US-Based credit card issuer, and **Measure & Monitor** the impact of the Dining Digest marketing campaign on customers across various metrics and dimensions

Problem Statement

Background

- The Member Services – Dining Team launched a quarterly e-Newsletter that informs card members of the benefits of using their card for dining spends

Problem Statement

- To setup the campaign with test and holdout populations that can be used to measure the quarter specific and overall impact.
- To measure the impact of the Dining Digest marketing campaign on customers across various metrics and dimensions
- To monitor the impact on a calendar month basis

Solution

- Setup of campaign with quarter specific and universal holdouts using randomizer
- Look at the change in metric from pre to post period for test and holdout population to measure impact of campaign
- Provide recommendations for the campaign design of subsequent eNLs

Approach

Methodology

Thought Leadership: Campaign Setup

Campaign was setup to measure impact at two levels

Metrics and Dimensions

Analysis Metrics

Overall and Dining Spend Metrics
- Spend ($)
- ROCs (#)
- Average Transaction Size ($)
- Spend Active %

Gold Dining Credit Metrics
- Dining Credit Redeemer %

Analysis Dimensions

Source Codes
Card Product
Tenure Bucket
Spend Active
Email Engagement
- Email Open
- Email Click

Impact Measurement

Impact per CM =

(post test metric – pre test metric)
–
(post control metric – pre control metric)

Impact | Example Insights

Per CM Impact

Show Statistical Siginifacnce No

Metric	April	May	June	July	August	Total
Spend per CM	$ 5.00	$ 2.00	$ (0.50)	$ (0.20)	$ 7.00	$ 13.30
Dining Spend per CM	$ 8.00	$ 3.00	$ 1.00	$ (2.00)	$ 4.00	$ 14.00
Dining Credit per CM	0.7	0.1	0	-0.3	0.2	$ 0.70

EXL

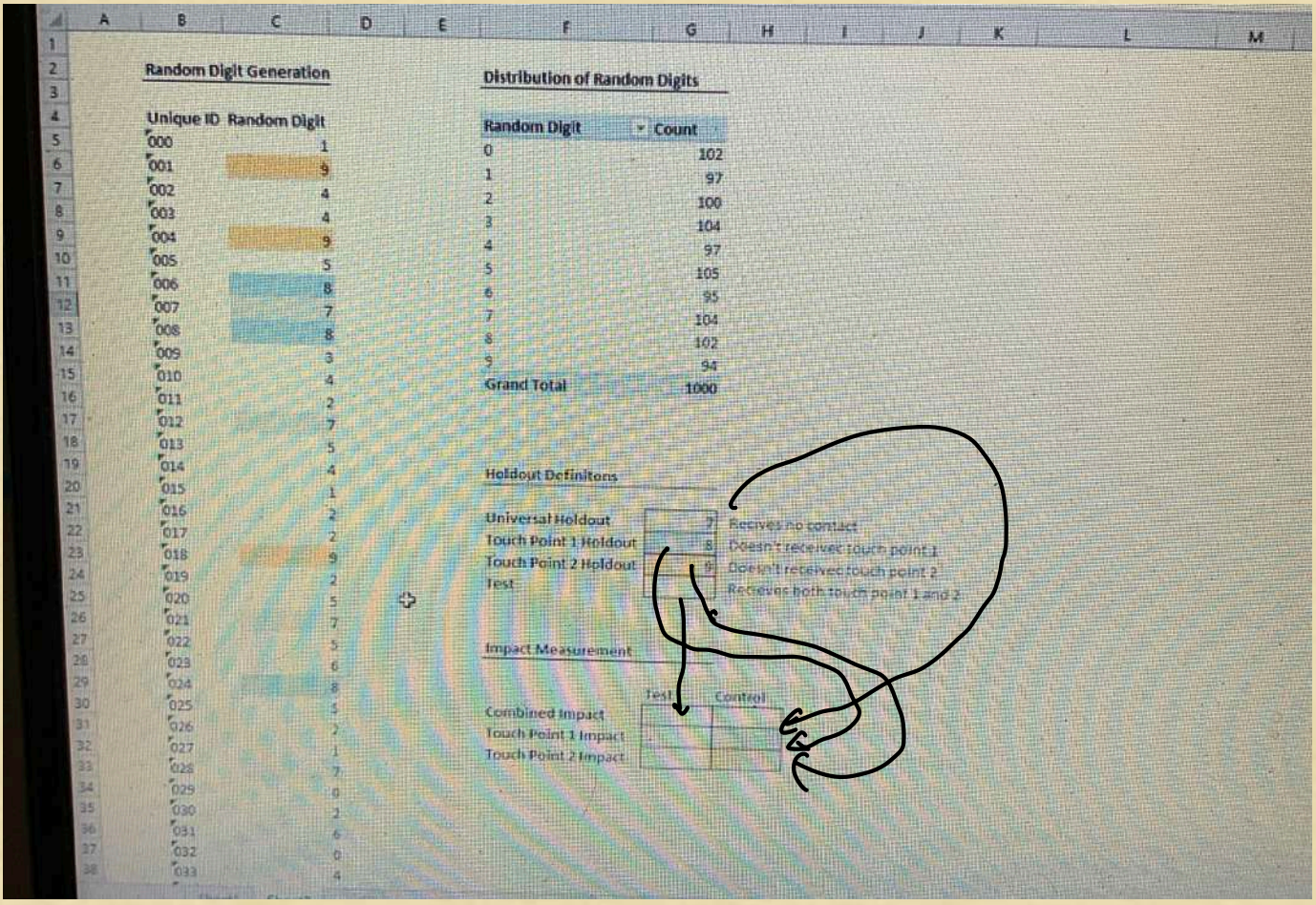
– Spend in Dining & overall level

Dashboard as self service tool

Cut & slice groups

– Cards Product
– Tenure Bucket
– Spend Active - 0 or 1
– Email - 0 or 1 opened

Random uniform sampling logic



Random Digit Generation

Unique ID	Random Digit
000	1
001	9
002	4
003	4
004	9
005	5
006	8
007	7
008	8
009	3
010	4
011	2
012	7
013	5
014	4
015	1
016	2
017	2
018	9
019	2
020	5
021	7
022	5
023	6
024	8
025	5
026	2
027	1
028	7
029	0
030	2
031	6
032	0
033	4

Distribution of Random Digits

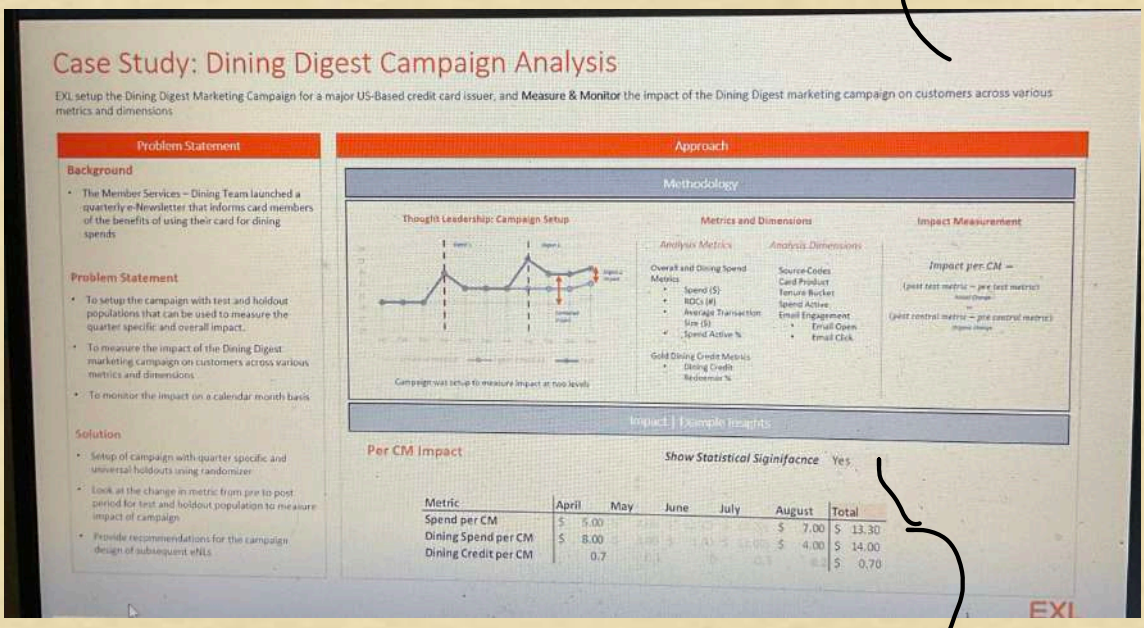
Random Digit	Count
0	102
1	97
2	100
3	104
4	97
5	105
6	95
7	104
8	102
9	94
Grand Total	1000

Holdout Definitons

Universal Holdout	7	Receives no contact
Touch Point 1 Holdout	8	Doesn't receives touch point 1
Touch Point 2 Holdout	9	Doesn't receives touch point 2
Test		Receives both touch point 1 and 2

Impact Measurement

	Test	Control
Combined Impact		
Touch Point 1 Impact		
Touch Point 2 Impact		



Case Study: Dining Digest Campaign Analysis

EXL setup the Dining Digest Marketing Campaign for a major US-Based credit card issuer, and **Measure & Monitor** the impact of the Dining Digest marketing campaign on customers across various metrics and dimensions

Problem Statement

Background

- The Member Services – Dining Team launched a quarterly e-Newsletter that informs card members of the benefits of using their card for dining spends

Problem Statement

- To setup the campaign with test and holdout populations that can be used to measure the quarter specific and overall impact.
- To measure the impact of the Dining Digest marketing campaign on customers across various metrics and dimensions
- To monitor the impact on a calendar month basis

Solution

- Setup of campaign with quarter specific and universal holdouts using randomizer
- Look at the change in metric from pre to post period for test and holdout population to measure impact of campaign
- Provide recommendations for the campaign design of subsequent eNLs

Approach

Methodology

Thought Leadership: Campaign Setup

Campaign was setup to measure impact at two levels

Metrics and Dimensions

Analysis Metrics

Overall and Dining Spend Metrics
- Spend ($)
- ROCs (#)
- Average Transaction Size ($)
- Spend Active %

Gold Dining Credit Metrics
- Dining Credit Redeemer %

Analysis Dimensions

Source-Codes
Card Product
Tenure Bucket
Spend Active
Email Engagement
- Email Open
- Email Click

Impact Measurement

Impact per CM =

(post test metric – pre test metric) / (post control metric – pre control metric)

Impact | Example Insights

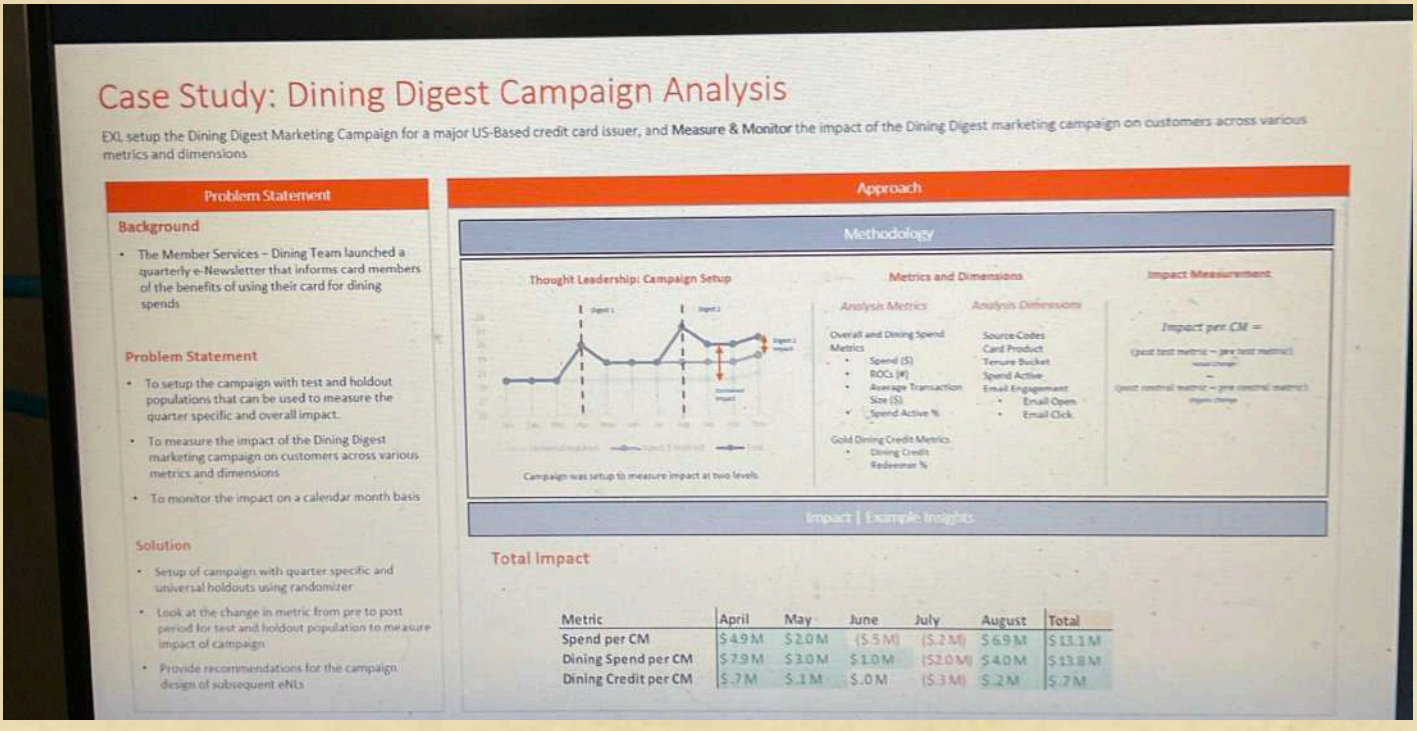
Per CM Impact

Show Statistical Siginifacnce Yes

Metric	April	May	June	July	August	Total
Spend per CM	$ 5.00				$ 7.00	$ 13.30
Dining Spend per CM	$ 8.00				$ 4.00	$ 14.00
Dining Credit per CM	0.7					$ 0.70

EXL

0.01 statistical significant.



Case Study: Dining Digest Campaign Analysis

EXL setup the Dining Digest Marketing Campaign for a major US-Based credit card issuer, and **Measure & Monitor** the impact of the Dining Digest marketing campaign on customers across various metrics and dimensions

Problem Statement

Background

- The Member Services – Dining Team launched a quarterly e-Newsletter that informs card members of the benefits of using their card for dining spends

Problem Statement

- To setup the campaign with test and holdout populations that can be used to measure the quarter specific and overall impact.
- To measure the impact of the Dining Digest marketing campaign on customers across various metrics and dimensions
- To monitor the impact on a calendar month basis

Solution

- Setup of campaign with quarter specific and universal holdouts using randomizer
- Look at the change in metric from pre to post period for test and holdout population to measure impact of campaign
- Provide recommendations for the campaign design of subsequent eNLs

Approach

Methodology

Thought Leadership: Campaign Setup

Campaign was setup to measure impact at two levels

Metrics and Dimensions

Analysis Metrics

Overall and Dining Spend Metrics
- Spend ($)
- ROCs (#)
- Average Transaction Size ($)
- Spend Active %

Gold Dining Credit Metrics
- Dining Credit Redeemer %

Analysis Dimensions

Source-Codes
Card Product
Tenure Bucket
Spend Active
Email Engagement
- Email Open
- Email Click

Impact Measurement

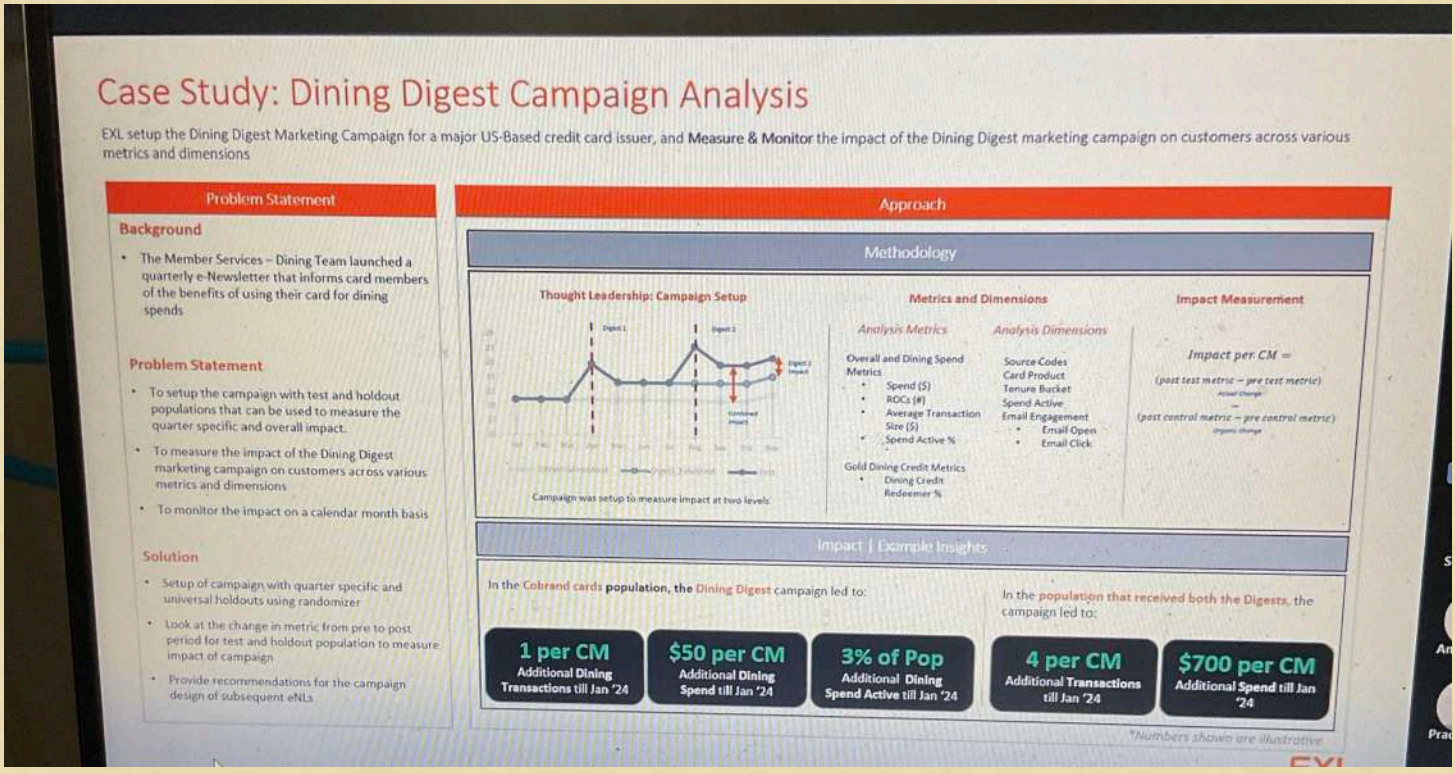
Impact per CM =

(post test metric – pre test metric) / (post control metric – pre control metric)

Impact | Example Insights

Total Impact

Metric	April	May	June	July	August	Total
Spend per CM	$4.9M	$2.0M	($.5M)	($.2M)	$6.9M	$13.1M
Dining Spend per CM	$7.9M	$3.0M	$1.0M	($2.0M)	$4.0M	$13.8M
Dining Credit per CM	$.7M	$.1M	$.0M	($.3M)	$.2M	$.7M



Sample was — 4 million.

- 5% Universal hold out.
- 5% Digest 1 hold out.
- 5% Digest 2 hold out

In each touchpoint it was always 10%

Digest 1 → 10% → 5% Universal hold out.
+ 5% Digest 1 hold out

Digest 2 → 10% → 5% Universal hold out
+ 5% Digest 2 hold out.

Power analysis for Test & Sampling

→ Industry → 50% Test
50% Control.

→ In Amex.

90% to 10%.
Control Test
↓
mostly 5% only.

→ 4M ? → eligibity.
→ dining rock
→ dining spend

only could decide 4M split was

$\delta = 5\%$?
size of =
test & Control
Power analysis
↓
min control size req.
1000 control size
↓
after campaign actually > 5% lift then no need of Statistical significance

In real life all information is only after campaign is run - → only after this we have to run the universal hold out group is re calculated if the pre campaign behaviour is different from test campaign behaviour