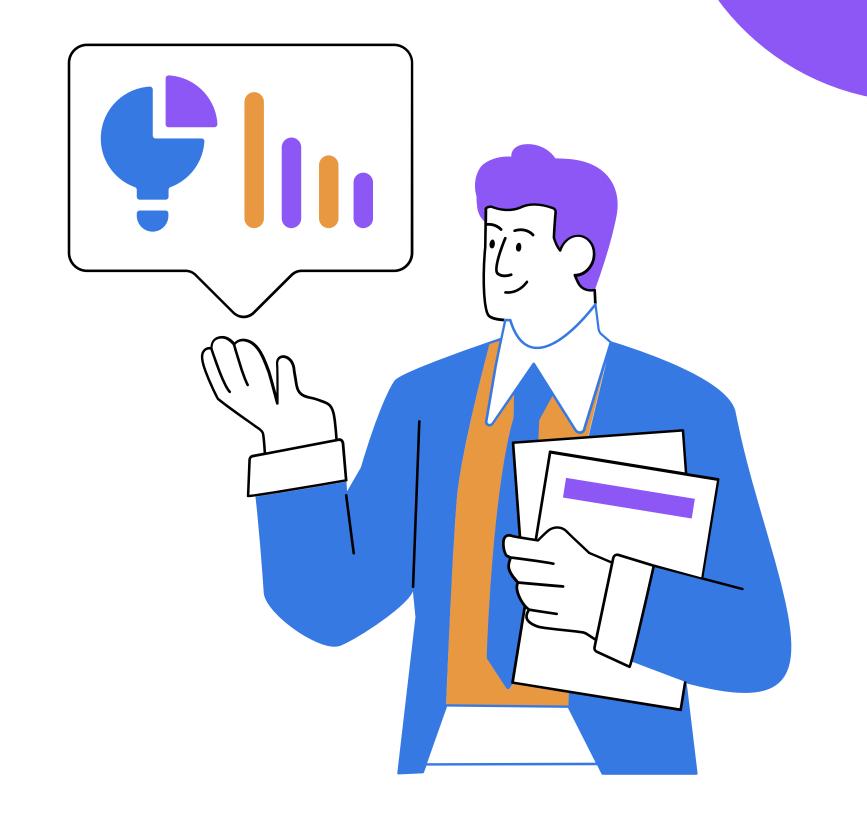
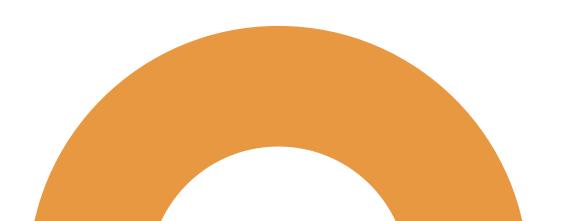
# ANALYSIS OF CREDIT CARD DEFAULTERS

Marwan Ahmed - 202101214





# Agenda

- Data Analysis
- Estimation Theory
- Confidence Intervals
- Hypothesis Testing/A-B Testing
- Regression Modeling

# WHICH CUSTOMERS ARE AT HIGHEST RISK OF DEFAULT?

### What is the overall distribution of key features?

#### Executives

Focus on trends and ROI from customer feedback.

#### Product managers

Use reviews to refine product features.

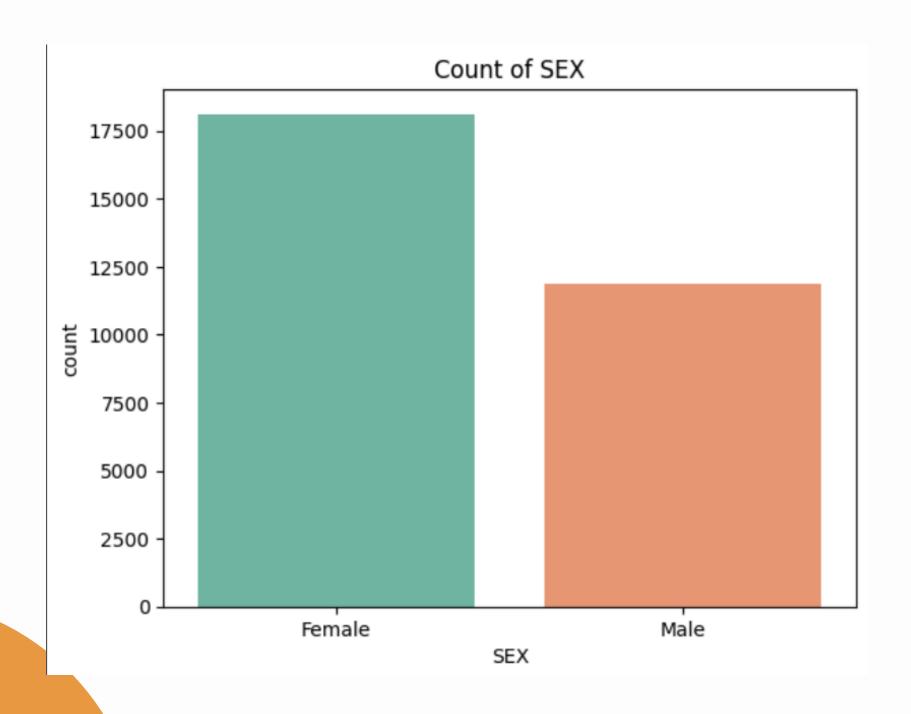
#### Marketing teams

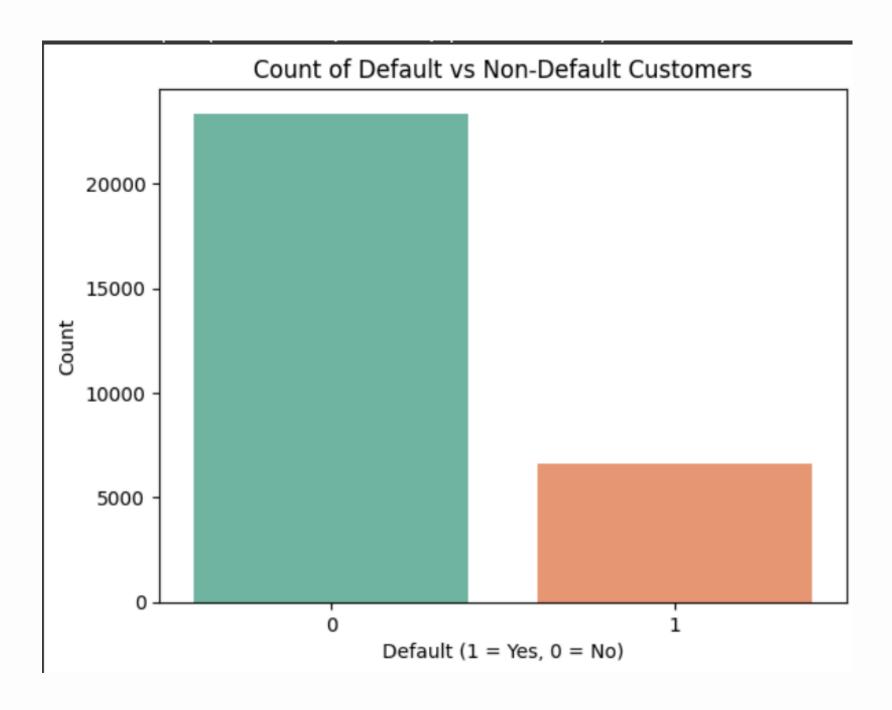
Monitor reputation and promote positive sentiment.

#### New Buyer

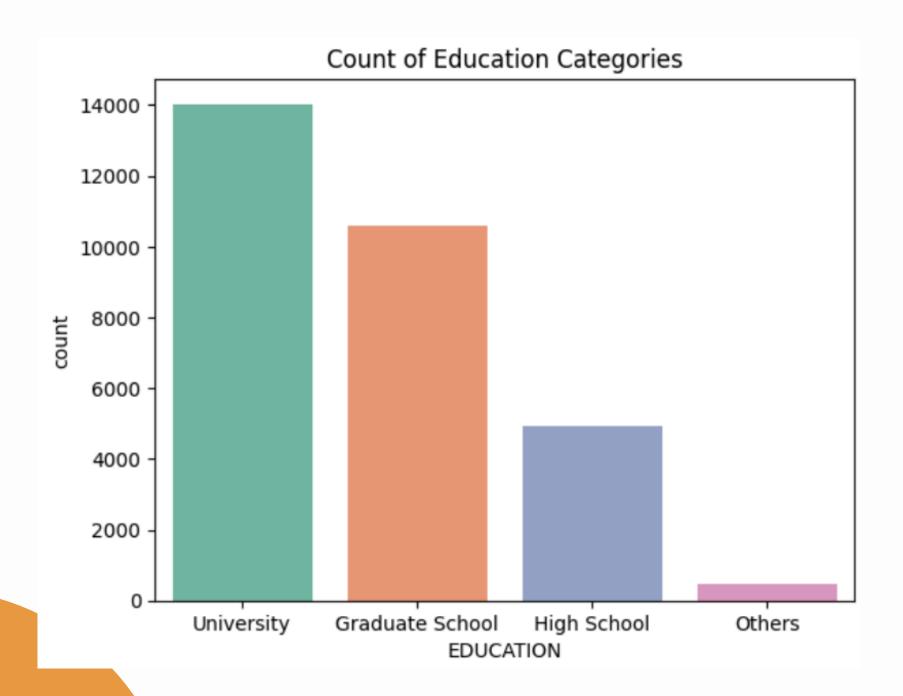
Relies on reviews before buying a product online

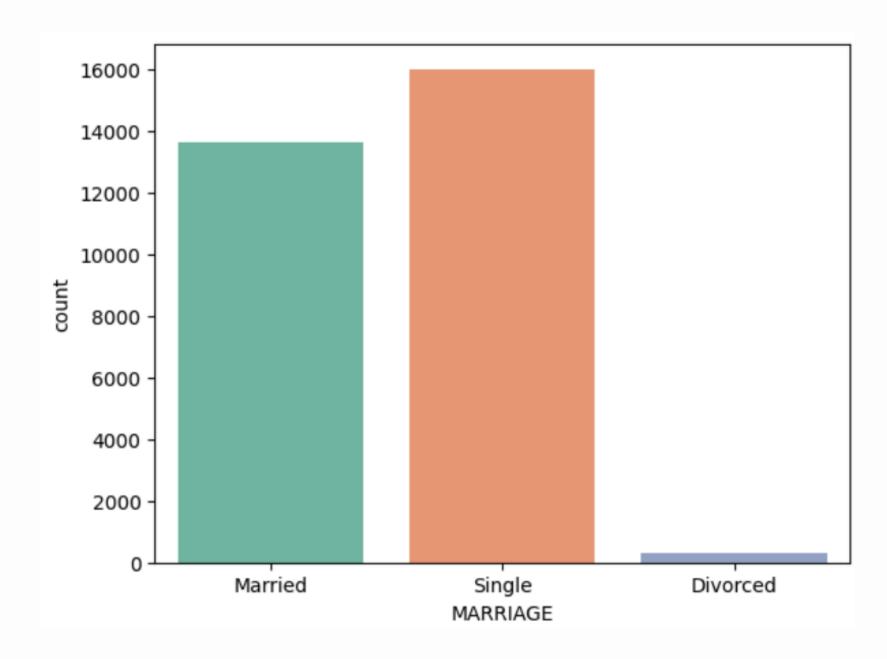
### What is the overall distribution of key features?

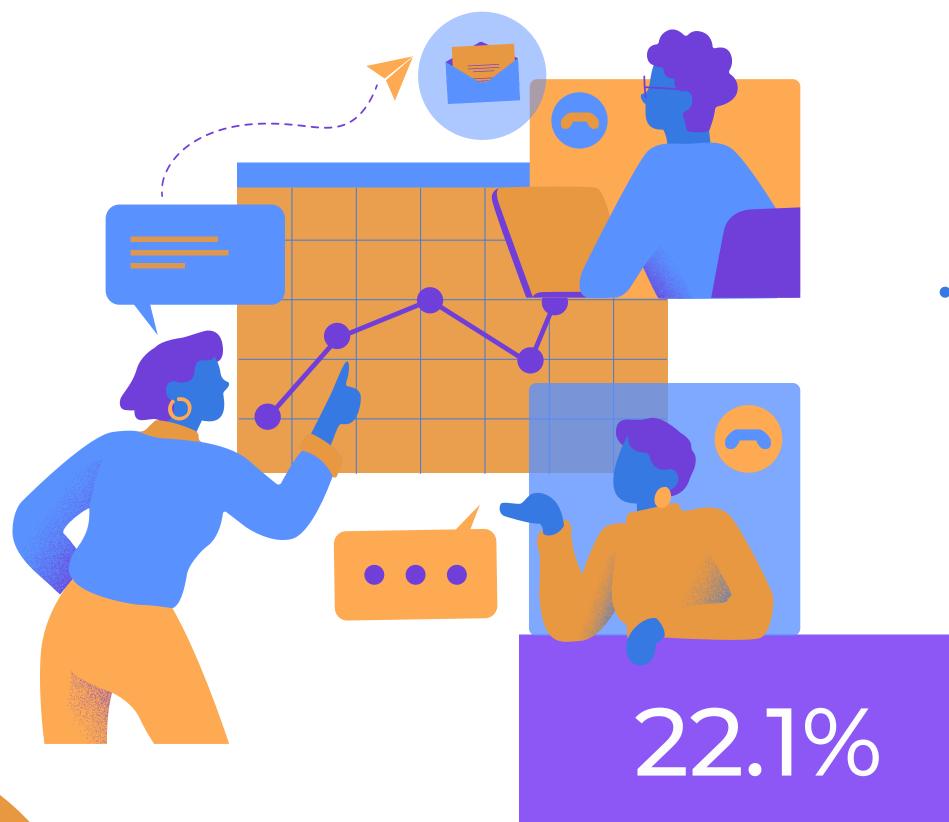




### What is the overall distribution of key features?

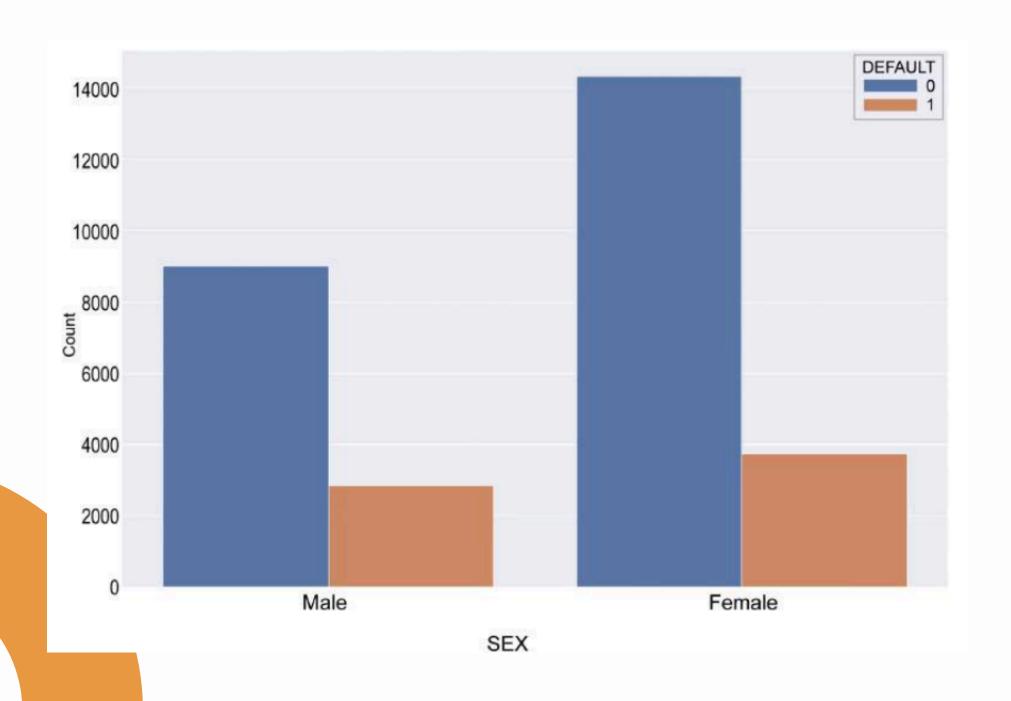






• 22.1% of customers default (DEFAULT countplot). Slightly more males than females, the largest education group is University, and most customers are single/married.

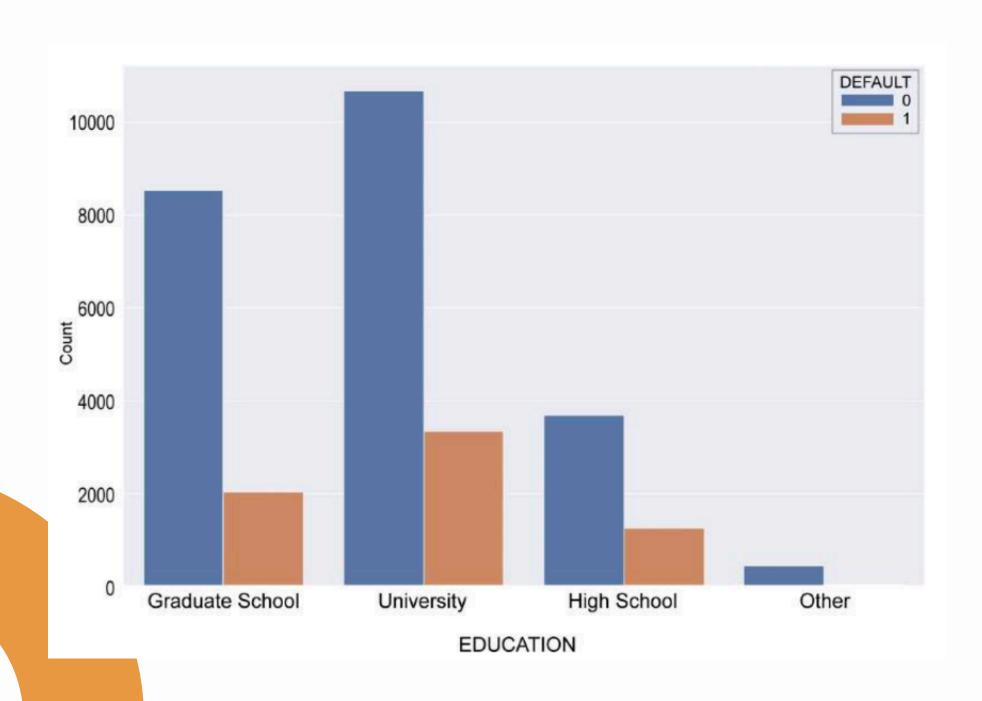
#### Do males default more than females?



24% of males default vs 20% of females.
 Males are at higher risk.

DEFAULT	0	1
SEX		
Female	0.792237	0.207763
Male	0.758328	0.241672
All	0.778800	0.221200

#### Does education level affect default?



 High-school graduates default ~25%, higher than University (~20%) or Graduate (~18%).

DEFAULT	0	1
EDUCATION		
University	10700	3330
Graduate School	8549	2036
High School	3680	1237
Others	435	33

#### Correlation of Features with DEFAULT

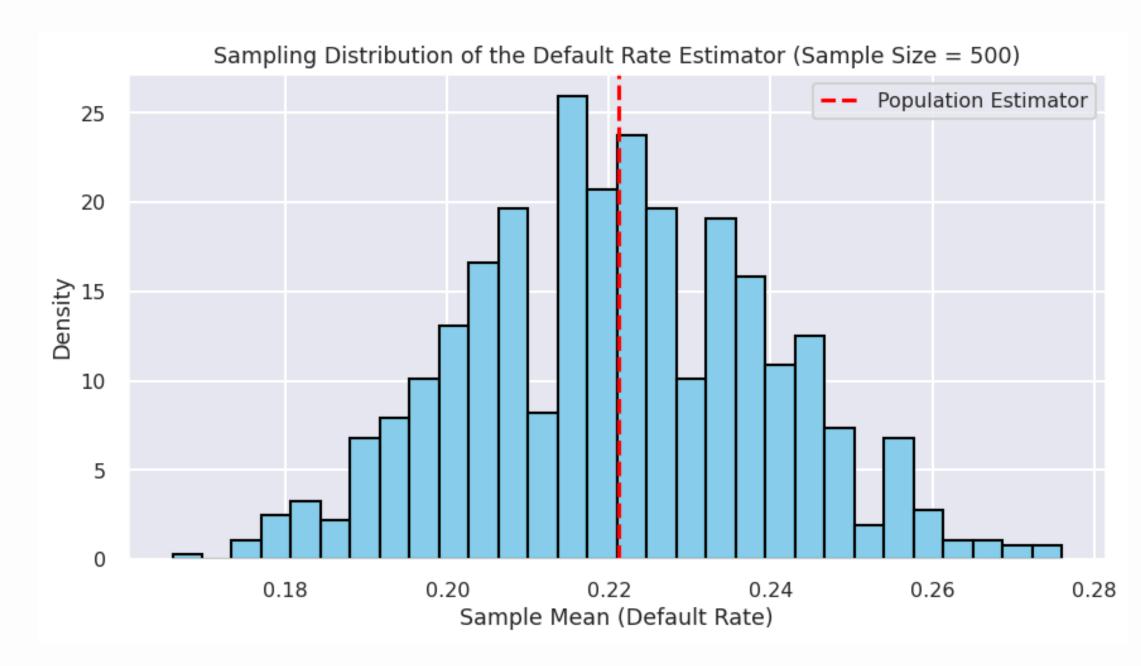
• We checked Spearman correlations. PAY\_1 (0.29) and PAY\_2 (0.22) are strongest predictors; PAYMENT AMT negatively correlated.

ID LIMIT BAL	1																					
I IMIT RAI		0.031	0.025	-0.025	-0.0049	-0.0092	-0.0043	-0.016	-0.0063	0.011	0.0097	0.016	0.033	0.015	0.02	0.012	0.051	0.094	0.022	0.012	0.038	-0.014
Dim_DAL	0.031	1	0.19	-0.3	-0.34	-0.33	-0.31	-0.28	-0.26	0.054	0.049	0.061	0.073	0.081	0.088	0.27	0.28	0.28	0.28	0.29	0.32	-0.17
AGE	0.025	0.19	1	-0.064	-0.083	-0.083	-0.08	-0.083	-0.076	0.001	0.0015	0.0018	-0.0033		0.00022	0.034	0.044	0.033	0.041	0.038	0.039	0.0051
PAY_1	-0.025	-0.3	-0.064	1	0.63	0.55	0.52	0.49	0.46	0.31	0.33	0.31	0.31	0.3	0.29	-0.098	-0.064	-0.054	-0.034	-0.026	-0.045	0.29
PAY_2	-0.0049	-0.34	-0.083	0.63	1	0.8	0.71	0.67	0.64	0.57	0.55	0.52	0.5	0.48	0.46	0.02	0.084	0.087	0.095	0.099	0.082	0.22
PAY_3	-0.0092	-0.33	-0.083	0.55	0.8	1	0.8	0.72	0.67	0.52	0.59	0.56	0.53	0.51	0.48	0.22	0.037	0.1	0.12	0.12	0.098	0.19
PAY_4	-0.0043	-0.31	-0.08	0.52	0.71	0.8	1	0.82	0.73	0.51	0.56	0.62	0.59	0.56	0.53	0.19	0.25	0.069	0.14	0.16	0.14	0.17
PAY_5	-0.016	-0.28	-0.083	0.49	0.67	0.72	0.82	1	0.82	0.5	0.54	0.59	0.65	0.62	0.58	0.18	0.22	0.26	0.11	0.18	0.17	0.16
PAY_6	-0.0063	-0.26	-0.076	0.46	0.64	0.67	0.73	0.82	1	0.49	0.52	0.56	0.61	0.67	0.63	0.18	0.2	0.24	0.28	0.14	0.2	0.14
BILL_AMT1	0.011	0.054	0.001	0.31	0.57	0.52	0.51	0.5	0.49	1	0.91	0.86	0.81	0.77	0.73	0.5	0.47	0.44	0.44	0.42	0.41	-0.025
BILL_AMT2	0.0097	0.049	0.0015	0.33	0.55	0.59	0.56	0.54	0.52	0.91	1	0.91	0.85	0.8	0.77	0.64	0.5	0.47	0.46	0.45	0.43	-0.016
BILL_AMT3	0.016	0.061	0.0018	0.31	0.52	0.56	0.62	0.59	0.56	0.86	0.91	1	0.9	0.85	0.8	0.55	0.64	0.49	0.49	0.48	0.46	-0.013
BILL_AMT4	0.033	0.073	-0.0033	0.31	0.5	0.53	0.59	0.65	0.61	0.81	0.85	0.9	1	0.9	0.85	0.51	0.55	0.63	0.51	0.5	0.48	-0.0084
BILL_AMT5	0.015	0.081	-0.00034	0.3	0.48	0.51	0.56	0.62	0.67	0.77	0.8	0.85	0.9	1	0.9	0.48	0.52	0.55	0.65	0.52	0.51	-0.0069
BILL_AMT6	0.02	0.088	0.00022	0.29	0.46	0.48	0.53	0.58	0.63	0.73	0.77	0.8	0.85	0.9	1	0.46	0.49	0.52	0.57	0.67	0.53	-7.6e-05
PAY_AMT1	0.012	0.27	0.034	-0.098	0.02	0.22	0.19	0.18	0.18	0.5	0.64	0.55	0.51	0.48	0.46	1	0.51	0.52	0.49	0.47	0.46	-0.16
PAY_AMT2	0.051	0.28	0.044	-0.064	0.084	0.037	0.25	0.22	0.2	0.47	0.5	0.64	0.55	0.52	0.49	0.51	1	0.52	0.52	0.5	0.49	-0.15
PAY_AMT3	0.094	0.28	0.033	-0.054	0.087	0.1	0.069	0.26	0.24	0.44	0.47	0.49	0.63	0.55	0.52	0.52	0.52	1	0.52	0.53	0.51	-0.14
PAY_AMT4	0.022	0.28	0.041	-0.034	0.095	0.12	0.14	0.11	0.28	0.44	0.46	0.49	0.51	0.65	0.57	0.49	0.52	0.52	1	0.53	0.55	-0.13
PAY_AMT5	0.012	0.29	0.038	-0.026	0.099	0.12	0.16	0.18	0.14	0.42	0.45	0.48	0.5	0.52	0.67	0.47	0.5	0.53	0.53	1	0.55	-0.12
PAY_AMT6	0.038	0.32	0.039	-0.045	0.082	0.098	0.14	0.17	0.2	0.41	0.43	0.46	0.48	0.51	0.53	0.46	0.49	0.51	0.55	0.55	1	-0.12
DEFAULT	-0.014	-0.17	0.0051	0.29	0.22	0.19	0.17	0.16	0.14	-0.025	-0.016	-0.013	-0.0084	-0.0069	-7.6e-05	-0.16	-0.15	-0.14	-0.13	-0.12	-0.12	1
	_		ш	_	٥.		4.	٠,	10	_	01	m	4	10	10		0	m	4	10	10	_
	Ω	BAL	AGE	PAY_1	PAY	PAY_3	PAY 4	PAY.	PAY_6	AMT	AMT2	AMT	AMT4	AMT5	AMT6	AMT	AMT2	AMT	AMT.	_AMT5	AMT(	DEFAULT
		MI				_	_	_		PILL_	BILL_	BILL_	BILL_	BILL_	BILL_	PAY_	PAY /	PAY_	PAY_	PAY_	PAY_AMT6	DEF

# How reliable is our default-rate estimate?

### Sampling Distribution of Default Rate

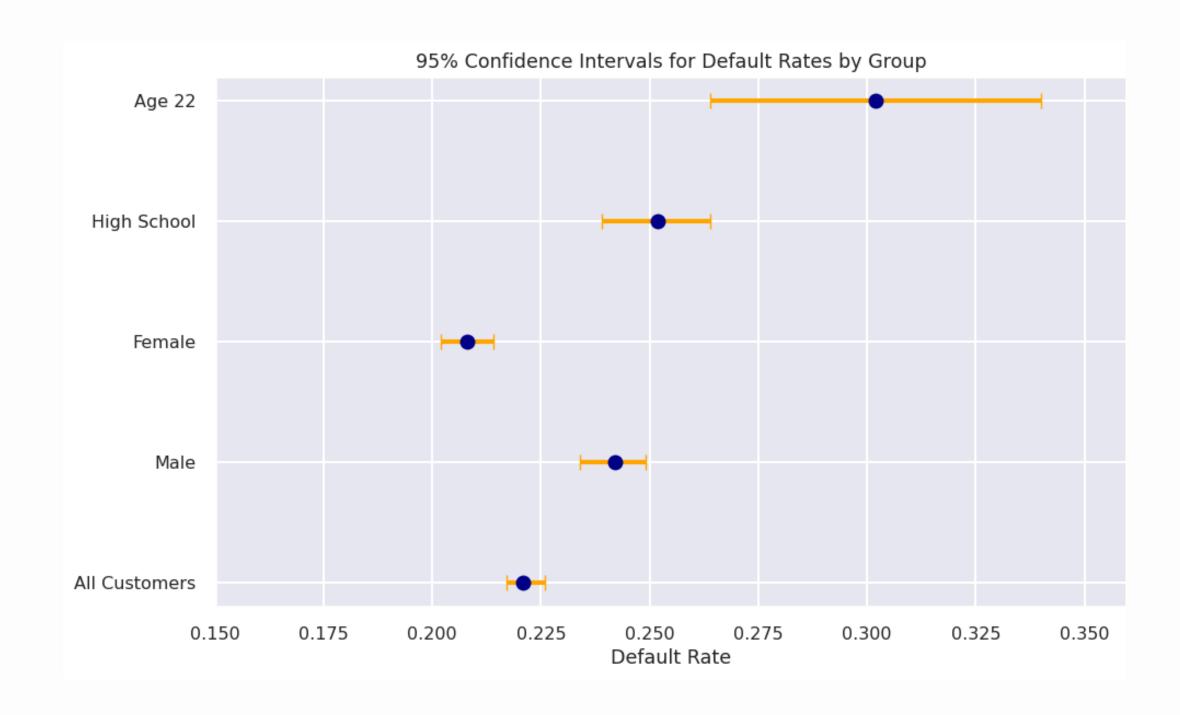
• We simulated 1,000 samples of 500 customers. The distribution of sample means is centered at ~22.2% with low variance, indicating our estimator is unbiased and stable.



# WHAT IS THE UNCERTAINTY AROUND OUR ESTIMATES?

#### 95% CIs for Default Rates

• All customers: 21.7%–22.6%. Males: 23.4–24.9%. Females: 20.2–21.4%. High School: 23.9–26.4%. Age 22: 26.4–34.0%.



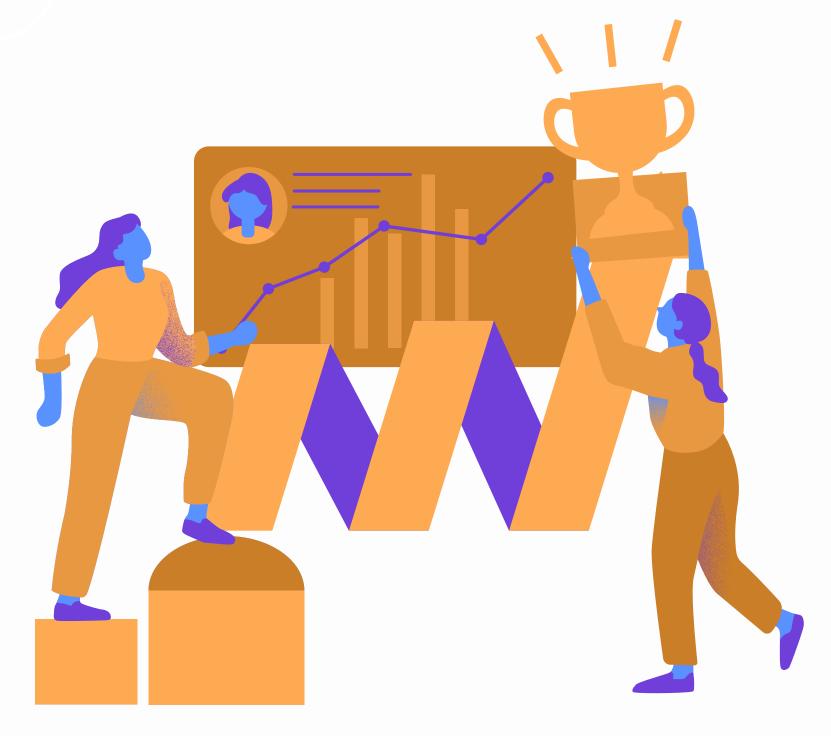
# ARE OBSERVED DIFFERENCES SIGNIFICANT?

## **Z-Tests for Key Comparisons**

Males vs females: Z=6.92, p<0.001 →
significant. HS vs Univ: Z=2.01, p=0.045 →
borderline significant. Age 22 vs others:
Z=4.64, p<0.001 → significant.</li>

## Conclusion & Actions

 High-risk customers: males, high-school educated, age 22, repeated late payers. Our combined EDA & inference workflow validates these insights.



####

# Thank Mountain

