

The background features a gradient from light green at the top to dark blue at the bottom. On the left side, there are several concentric circular patterns and a large arc with a scale ranging from 140 to 260. The scale is marked with numbers every 10 units. The overall design is technical and futuristic.

# LAB 5: PEER-TO-PEER LENDING

VINIT SHAH

# PART 2

- 41.61% of the loans are "HighGrade"
- Debtor Above or Below Media Income Level

## Welch Two Sample t-test

```
data: train$highgrade by highincome
t = -45.042, df = 231180, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.09534142 -0.08738990
sample estimates:
mean in group FALSE mean in group TRUE
      0.3728698      0.4642354
```

- Loan request is above or below the median loan amount

## Welch Two Sample t-test

```
data: train$highgrade by highloanamount
t = 34.091, df = 235540, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.06508972 0.07303066
sample estimates:
mean in group FALSE mean in group TRUE
      0.4505373      0.3814771
```



# PART 2

- Debtor rents their home or not

## Welch Two Sample t-test

```
data: train$highgrade by renters
t = 14.688, df = 199440, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.02638399 0.03450975
sample estimates:
mean in group FALSE  mean in group TRUE
      0.4280667      0.3976199
```

# PART 3

- Model

```
Call:
glm(formula = train$highgrade ~ train$annual_inc + train$home_ownership +
    train$loan_amnt, family = binomial)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-8.4904	-1.0284	-0.8749	1.2716	1.7719

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	7.522e+00	2.666e+01	0.282	0.778
train\$annual_inc	7.281e-06	1.149e-07	63.376	<2e-16 ***
train\$home_ownershipMORTGAGE	-7.699e+00	2.666e+01	-0.289	0.773
train\$home_ownershipOWN	-7.774e+00	2.666e+01	-0.292	0.771
train\$home_ownershipRENT	-7.853e+00	2.666e+01	-0.294	0.768
train\$loan_amnt	-4.331e-05	5.959e-07	-72.671	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 319984 on 235628 degrees of freedom  
Residual deviance: 312626 on 235623 degrees of freedom  
AIC: 312638

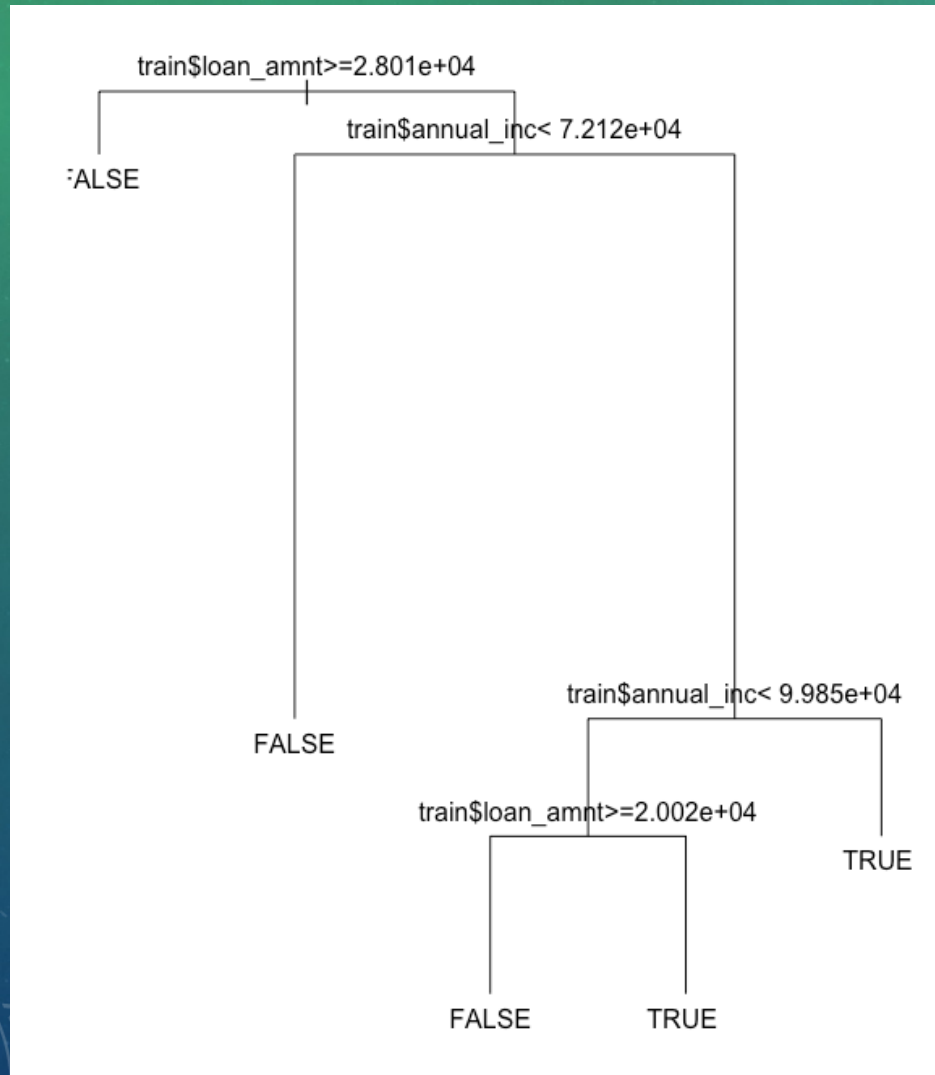
Number of Fisher Scoring iterations: 6

classi		
	FALSE	TRUE
FALSE	122809	14777
TRUE	80714	17329



# PART 4

- Classification Tree



aa			
	FALSE	TRUE	
FALSE	107469	30117	
TRUE	61930	36113	

# PART 5

lmodel_classified			
highgrade	FALSE	TRUE	
FALSE	181745	48408	
TRUE	122227	68715	

treemodel_classified			
highgrade	FALSE	TRUE	
FALSE	135029	95124	
TRUE	65704	125238	