

logistic Reg in r:

code:

```
# Load the necessary library
```

```
library(ggplot2)
```

```
# Load the data
```

```
data <- diamonds
```

```
# Create a binary outcome variable
```

```
data$expensive <- ifelse(data$price > 5000, 1, 0)
```

```
# Check the correlation between carat and the new binary variable
```

```
cor(data$carat, data$expensive)
```

```
# Plot the data
```

```
ggplot(data, aes(x = carat, y = expensive)) +
```

```
  geom_point() +
```

```
  labs(
```

```
    y = "Expensive (1 = Yes, 0 = No)",
```

```
    x = "Carat"
```

```
  ) +
```

```
  theme_minimal()
```

```
# Logistic regression
```

```
model <- glm(expensive ~ carat, family = binomial, data = data)
```

```
summary(model)
```

output:

```
> summary(model)
```

```
Call:
```

```
glm(formula = expensive ~ carat, family = binomial, data = data)
```

```
Coefficients:
```

	Estimate	Std. Error	z value	Pr(> z )	
(Intercept)	-12.6308	0.1561	-80.89	<2e-16	***
carat	12.0359	0.1510	79.70	<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: 63219  on 53939  degrees of freedom
```

```
Residual deviance: 18577  on 53938  degrees of freedom
```

```
AIC: 18581
```

```
Number of Fisher Scoring iterations: 8
```

