

DATA 621 Assignment 2

Kai Lukowiak

2018-03-14

Import Data

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.4.2
```

```
## Warning: package 'tibble' was built under R version 3.4.3
```

```
## Warning: package 'tidyr' was built under R version 3.4.3
```

```
## Warning: package 'purrr' was built under R version 3.4.2
```

```
## Warning: package 'dplyr' was built under R version 3.4.2
```

```
library(knitr)
```

```
df <- read_csv('~/.DATA621/Assignments/Assignment2/classification-output-data.csv')
```

```
sample_n(df, size = 5) %>% kable()
```

pregnant	glucose	diastolic	skinfold	insulin	bmi	pedigree	age	class	scored.class	scored.probability
1	90	68	8	0	24.5	1.138	36	0	0	0.1070070
0	147	85	54	0	42.8	0.375	24	0	0	0.4368037
3	113	50	10	85	29.5	0.626	25	0	0	0.1608049
1	124	60	32	0	35.8	0.514	21	0	0	0.2059407
0	105	90	0	0	29.6	0.197	46	0	0	0.1339198

```
x <- table(df$scored.class, df$class)
```

```
colnames(x) <- c('Actual Positive', 'Actual Negative')
```

```
rownames(x) <- c('Predicted Positive', 'Predicted Negative')
```

```
x %>% kable()
```

	Actual Positive	Actual Negative
Predicted Positive	119	30
Predicted Negative	5	27