Tutorial 3 – Lab Exercise Report

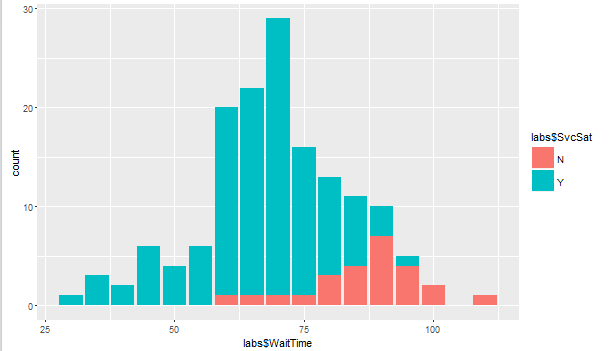
Chen Lu, A0105483

**c) Is there any statistical evidence to suggest that service satisfaction is correlated with waiting times?**

Answer:

#correlation between SvcSat & WaitTime

ggplot(data=labs, aes(x=labs$WaitTime, fill=labs$SvcSat)) + geom\_bar(stat= "count")



When the waiting time increases, more users tend to be unsatisfied with the service.

**d) Compute the Body Mass Index for each patient and test whether the ‘BMIs’ are correlated with ‘Total Cholesterol’**

Code:

#calculate BMI

labs$BMI <- (labs$WT / (labs$HT^2))

#check BMI & TC correlation

lm <- lm(labs$BMI ~ labs$TC)

summary(lm)

Output:

Call:

lm(formula = labs$BMI ~ labs$TC)

Residuals:

Min 1Q Median 3Q Max

-8.2196 -2.6349 -0.2814 2.3899 9.4007

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 19.21560 0.57056 33.678 < 2e-16 \*\*\*

labs$TC 0.06659 0.00980 6.795 2.43e-10 \*\*\*

---

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

Residual standard error: 3.4 on 149 degrees of freedom

Multiple R-squared: 0.2366, Adjusted R-squared: 0.2314

F-statistic: 46.17 on 1 and 149 DF, p-value: 2.431e-10

The high R-squared value and small p-value suggests there is a correlation between patients’ ‘BMIs’ and ‘Total Cholesterol’.

**e) Predict the ‘BMI’ of a Chinese male patient whose age is 45 and SBP is 135 and TC is 150. What assumptions are necessary for the prediction to be valid?**

Code:

group\_by(labs, Gender="M", ETHNIC = "C")

labs$TC <- as.numeric(labs$TC)

fit <- lm(BMI ~ AGE + SBP + TC, data = labs)

summary(fit)

newdata = data.frame(AGE = 45, SBP = 135, TC = 150)

prediction <- predict(fit, newdata)

prediction

Output:

> newdata = data.frame(AGE = 45, SBP = 135, TC = 150)

> prediction <- predict(fit, newdata)

> prediction

1

29.78424