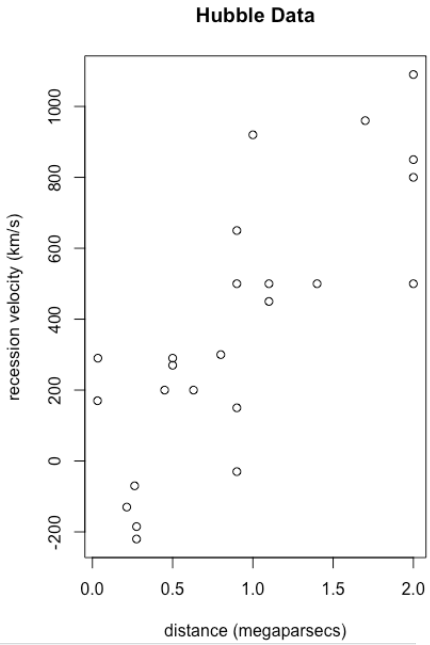
Luke Beebe
Hubble Data Assignment 2

1) The plot is posted to the right.
2) Yes, it would be appropriate. It looks like the relationship between distance and recession velocity is a positive linear relationship.
3) As the distance increases, the recession velocity increases at a rate of 454.16 km/s.
4) The data suggests there is a statistically significant relationship between the two variables. The p-value of the explanatory variable is 4.48e-06, below alpha=0.001.
5) Distance explains velocity 60.64% of the time, our R^2 value.
6) An estimate of Andromeda's recession velocity is 308.9184 km/s. I'm 95% certain its actual recession velocity is between (-184.5624, 802.3992)




```
#Regression Methods Assignment 2

hubble_data<-read.table("/Users/lukebeebe/Documents/School/Rutgers/Spring 2023/Regression Methods/hubble_data.txt")
colnames(hubble_data)<-c("distance","recession_velocity")
hubble_data<-hubble_data[2:25,]
hubble_data #1 below
plot(hubble_data$distance,hubble_data$recession_velocity,xlab="distance (megaparsecs)",ylab="recession velocity (km/s)",main="Hubble Data")
#2 Yes, it would be appropriate. From an eye inspection, it looks as if there is a positive linear relationship between distance and recession velocity.
hubble_data$distance<-as.numeric(as.character(hubble_data$distance))
hubble_model<-lm(recession_velocity~distance,data=hubble_data)
hubble_model #3 As the distance increases, the recession velocity increases at a rate of 454.16 km/s
summary(hubble_model) #4 The data suggests there is a statistically significant relationship between the two variable. The p-value of the explanatory variab
#5 Distance explains recession velocity 60.64% of the time, our R^2 value.
predict(hubble_model,data.frame(distance=0.77),interval = "predict",level = 0.95)
#6 An estimate of Andromeda's recession velocity is 308.9184 km/s, and I'm 95% certain it's actual recession velocity is between (-184.5624, 802.3992)
```

Above is my code, and to the right is the output of my code, where my answers were derived from.

```
> hubble_model #3 As the distance increases, the recession velocity increases at a r

Call:
lm(formula = recession_velocity ~ distance, data = hubble_data)

Coefficients:
(Intercept)     distance
     -40.78       454.16

> summary(hubble_model) #4 The data suggests there is a statistically significant re
s 4.48e-06, below alpha=0.001

Call:
lm(formula = recession_velocity ~ distance, data = hubble_data)

Residuals:
    Min      1Q  Median      3Q     Max
-397.96 -158.10  -13.16  148.09  506.63

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -40.78      83.44  -0.489     0.63
distance      454.16      75.24   6.036 4.48e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 232.9 on 22 degrees of freedom
Multiple R-squared:  0.6235,    Adjusted R-squared:  0.6064
F-statistic: 36.44 on 1 and 22 DF,  p-value: 4.477e-06

> #5 Distance explains recession velocity 60.64% of the time, our R^2 value.
> predict(hubble_model,data.frame(distance=0.77),interval = "predict",level = 0.95)
       fit       lwr      upr
1 308.9184 -184.5625 802.3992
```