Ben DuPey / Homework 1 / STAT341-A

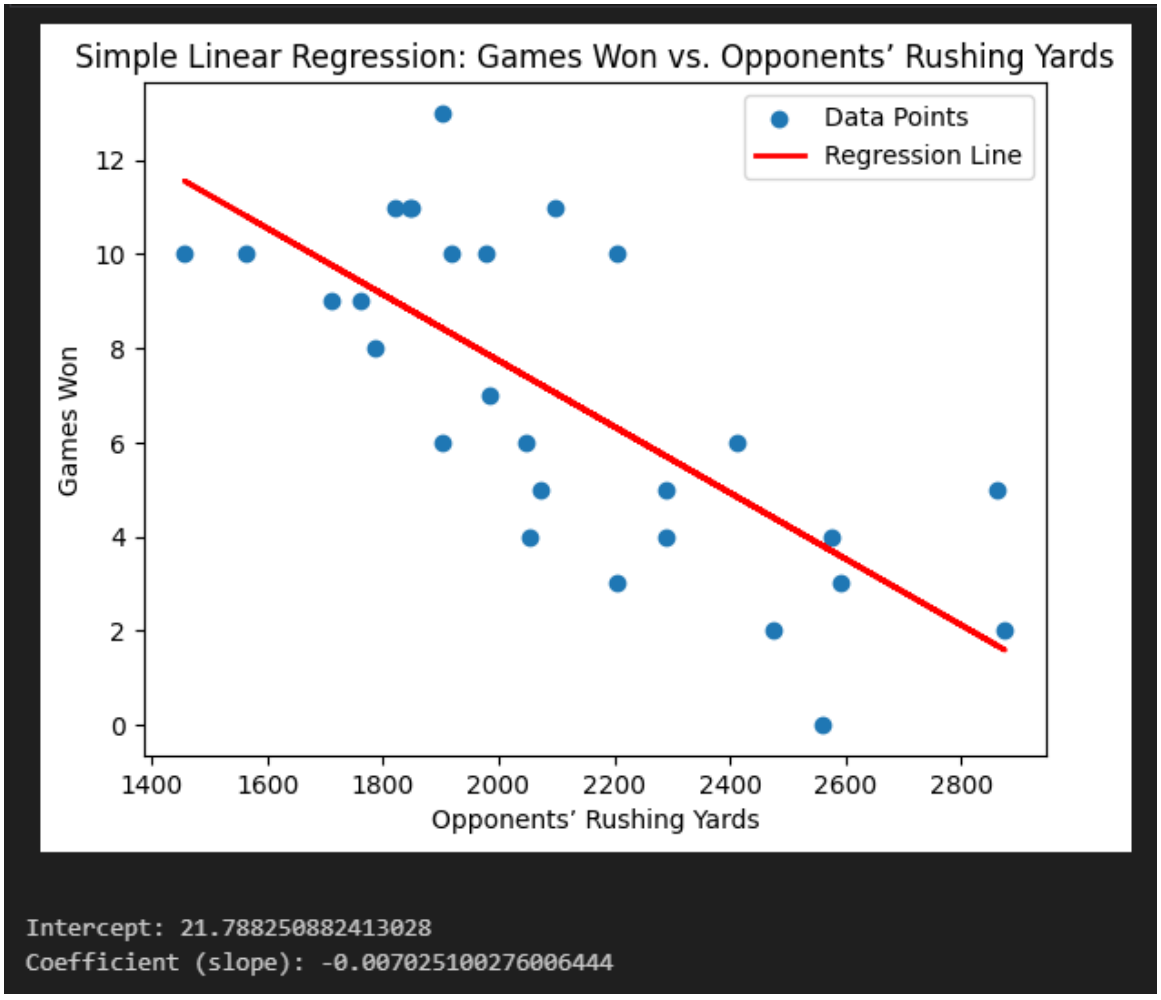
ATTENTION!! ALL OF MY CODE CAN BE VIEWED AT THIS LINK:

https://github.com/bdupey/STAT-341A_Appplied_Regression_Analysis_Modeling/blob/main/homework_1.ipynb

ALL QUESTIONS ARE CLEARLY LABELED.

ALL OF MY SOLUTIONS ARE INCLUDED IN THIS DOCUMENT.

2.1.a)



(Continued…)

2.1.B)

OLS Regression Results

==

Dep. Variable: y R-squared: 0.545

Model: OLS Adj. R-squared: 0.527

Method: Least Squares F-statistic: 31.10

Date: Thu, 14 Sep 2023 Prob (F-statistic): 7.38e-06

Time: 12:45:07 Log-Likelihood: -63.123

No. Observations: 28 AIC: 130.2

Df Residuals: 26 BIC: 132.9

Df Model: 1

Covariance Type: nonrobust

==

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	21.7883	2.696	8.081	0.000	16.246	27.330
x8	-0.0070	0.001	-5.577	0.000	-0.010	-0.004

==

Omnibus: 2.076 Durbin-Watson: 1.566

Prob(Omnibus): 0.354 Jarque-Bera (JB): 1.402

Skew: 0.305 Prob(JB): 0.496

Kurtosis: 2.089 Cond. No. 1.28e+04

2.1.C)

95% Confidence Interval for the Slope of x8:

0 -0.009614

1 -0.004436

Name: x8, dtype: float64

2.1.D)

	0	1
Intercept	16.246064	27.330438
x8	-0.009614	-0.004436
95% CI =	[-0.009614,	-0.004436]

2.1.E)

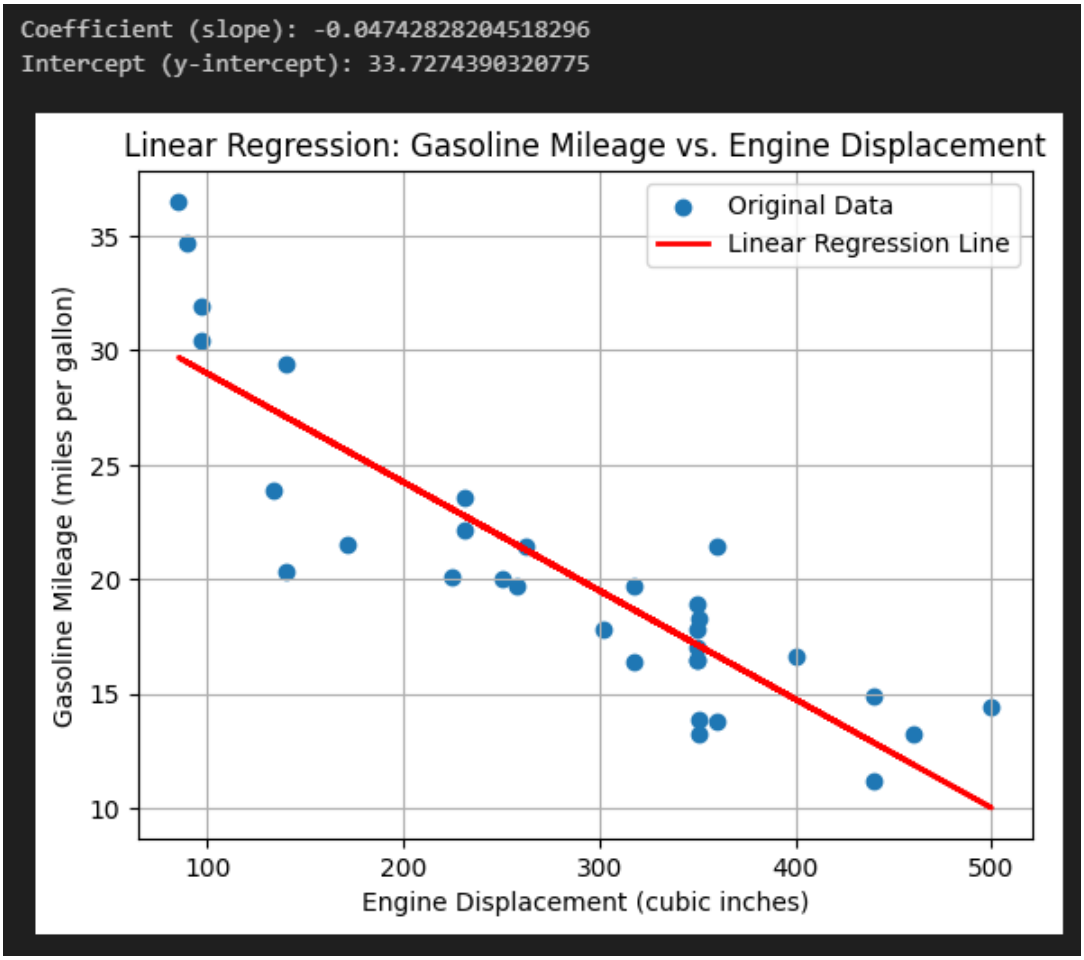
95% CI on the mean number of games won when x8 = 2000 yards: [2.72, 12.75]

2.2)

Point estimate of games won when x8 = 1800 yards: 9.14

90% Prediction interval on games won: [4.94, 13.35]

2.4.A)



2.4.B)

OLS Regression Results

Dep. Variable:	y	R-squared:	0.772
Model:	OLS	Adj. R-squared:	0.764
Method:	Least Squares	F-statistic:	101.6
Date:	Sat, 16 Sep 2023	Prob (F-statistic):	3.82e-11
Time:	12:28:37	Log-Likelihood:	-80.236
No. Observations:	32	AIC:	164.5
Df Residuals:	30	BIC:	167.4
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	33.7274	1.446	23.332	0.000	30.775	36.680
x1	-0.0474	0.005	-10.078	0.000	-0.057	-0.038

Omnibus:	0.208	Durbin-Watson:	1.668
Prob(Omnibus):	0.901	Jarque-Bera (JB):	0.152
Skew:	0.145	Prob(JB):	0.927
Kurtosis:	2.825	Cond. No.	819.

2.4.C)

Percent of Total Variability Explained: 77.20%

2.4.D)

95% Confidence Interval on Mean Gasoline Mileage (when x1 = 275 cubic inches):

Lower Bound: 14.42 miles per gallon

Upper Bound: 26.95 miles per gallon

2.4.E)

Point Estimate of Mileage (when x1 = 275 cubic inches): 20.68 miles per gallon

95% Prediction Interval on Mileage:

Lower Bound: 19.57 miles per gallon

Upper Bound: 21.80 miles per gallon

2.4.F)

The difference is that the confidence interval is estimating the range of the population mean (average mileage for all cars with a 275-in.^{3} engine), which can be quite wide due to sample variability. In contrast, the prediction interval is estimating the range of a single observation (the mileage of one specific car with that engine size), and it is narrower because it is focused on predicting a specific value rather than summarizing a population parameter.

2.7.A)

Coefficient (slope): 0.03297360191329696

Intercept (y-intercept): -1.8450701804751002

2.7.B)

T-Statistic: 3.3861

P-Value: 0.0033

Reject the null hypothesis ($H_0: \beta_1 = 0$)

2.7.C)

R-squared: 0.3891

2.7.D)

95% Confidence Interval on the Slope (beta1):

Lower Bound: 0.0125

Upper Bound: 0.0534

2.7.E)

Predicted Mean Purity: 89.66431235475011

95% Confidence Interval: (array([88.16896658]), array([91.15965813]))