8th Grade

DS106-01-11

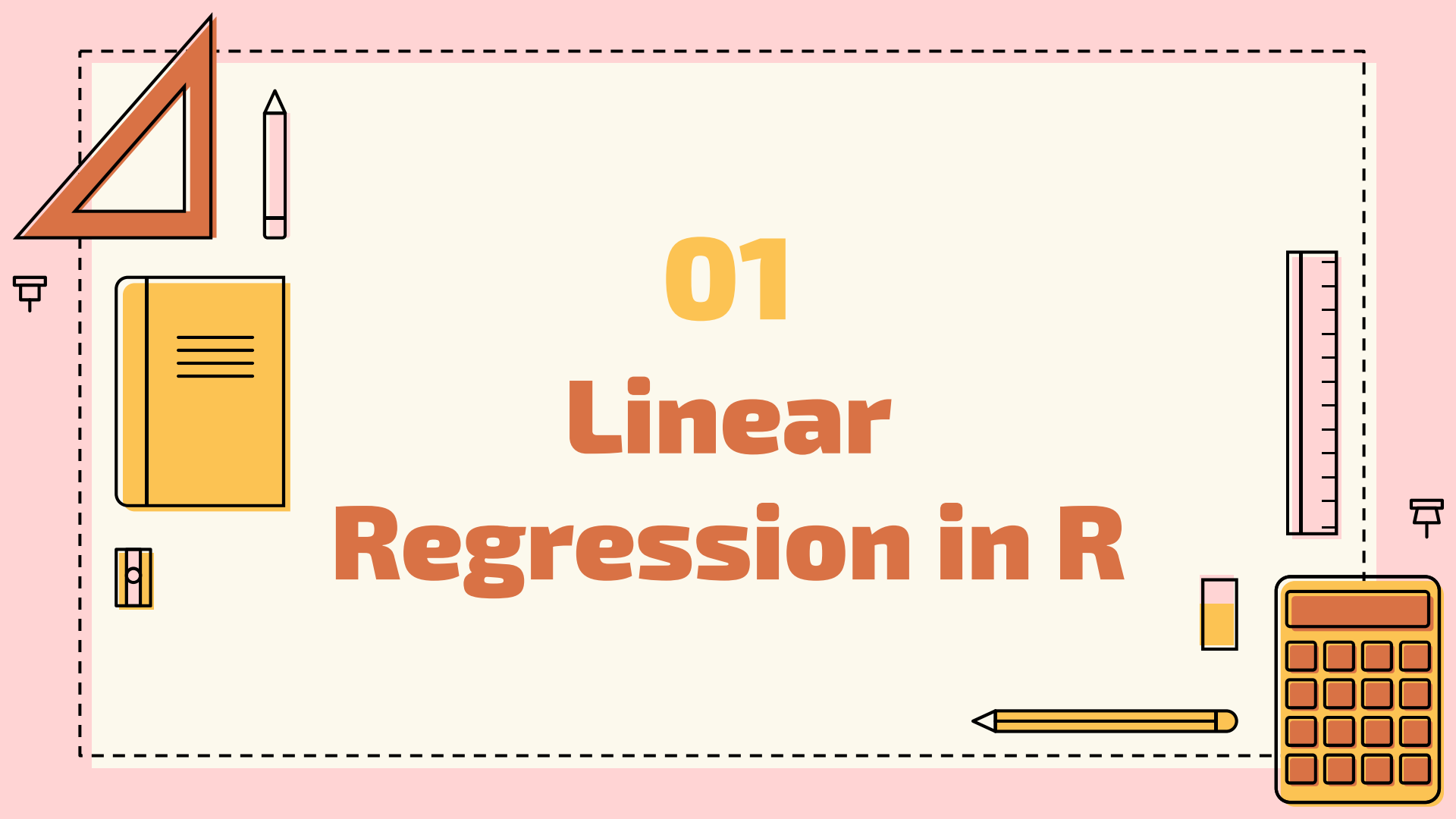
Lesson 1 Practice Hands-On

Heather Walker - 2022-10-26

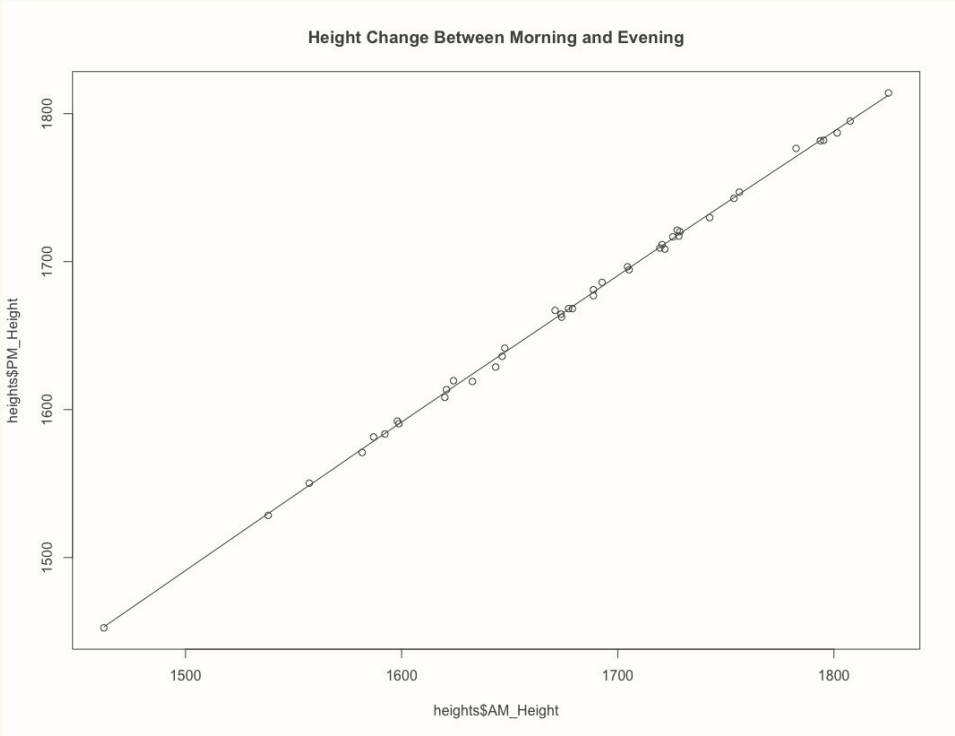
Table of contents

01

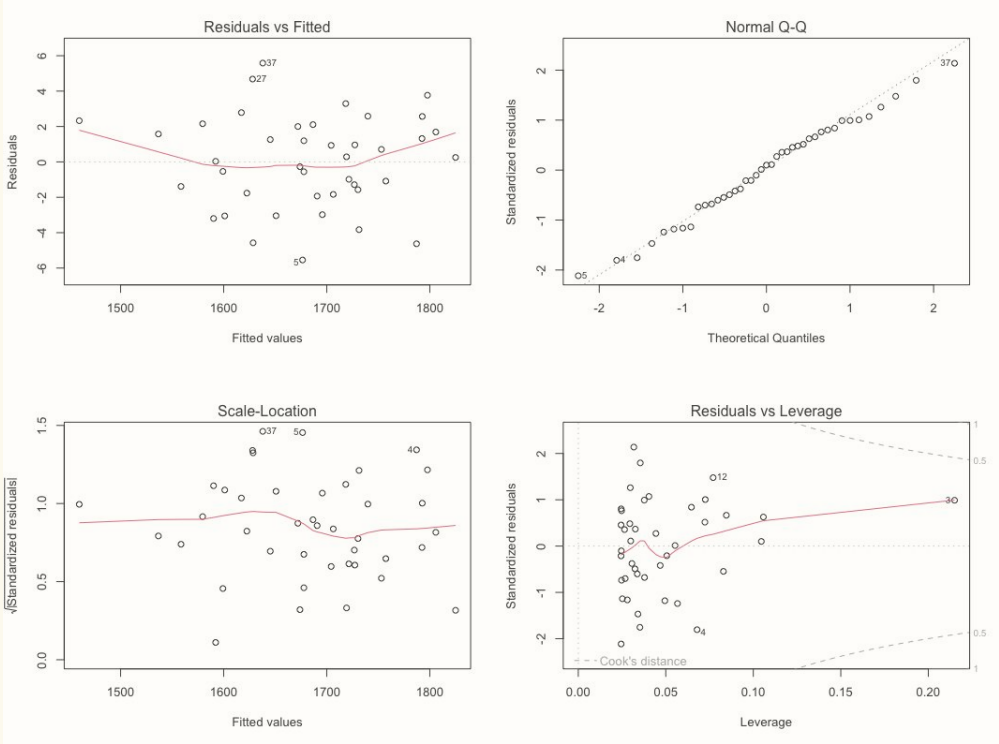
Linear Regression in R



Testing for Linearity



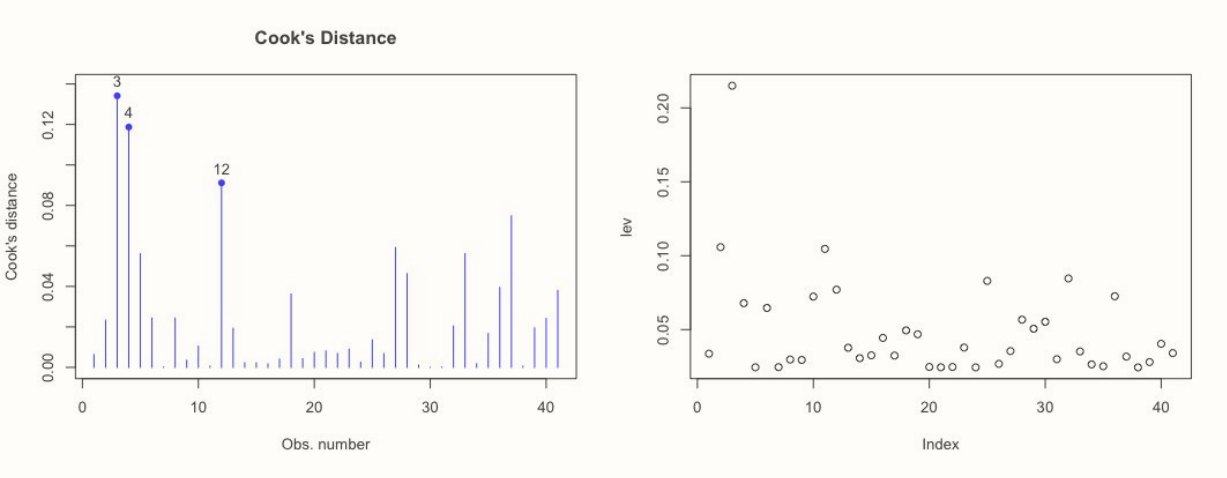
Testing for Homoscedasticity



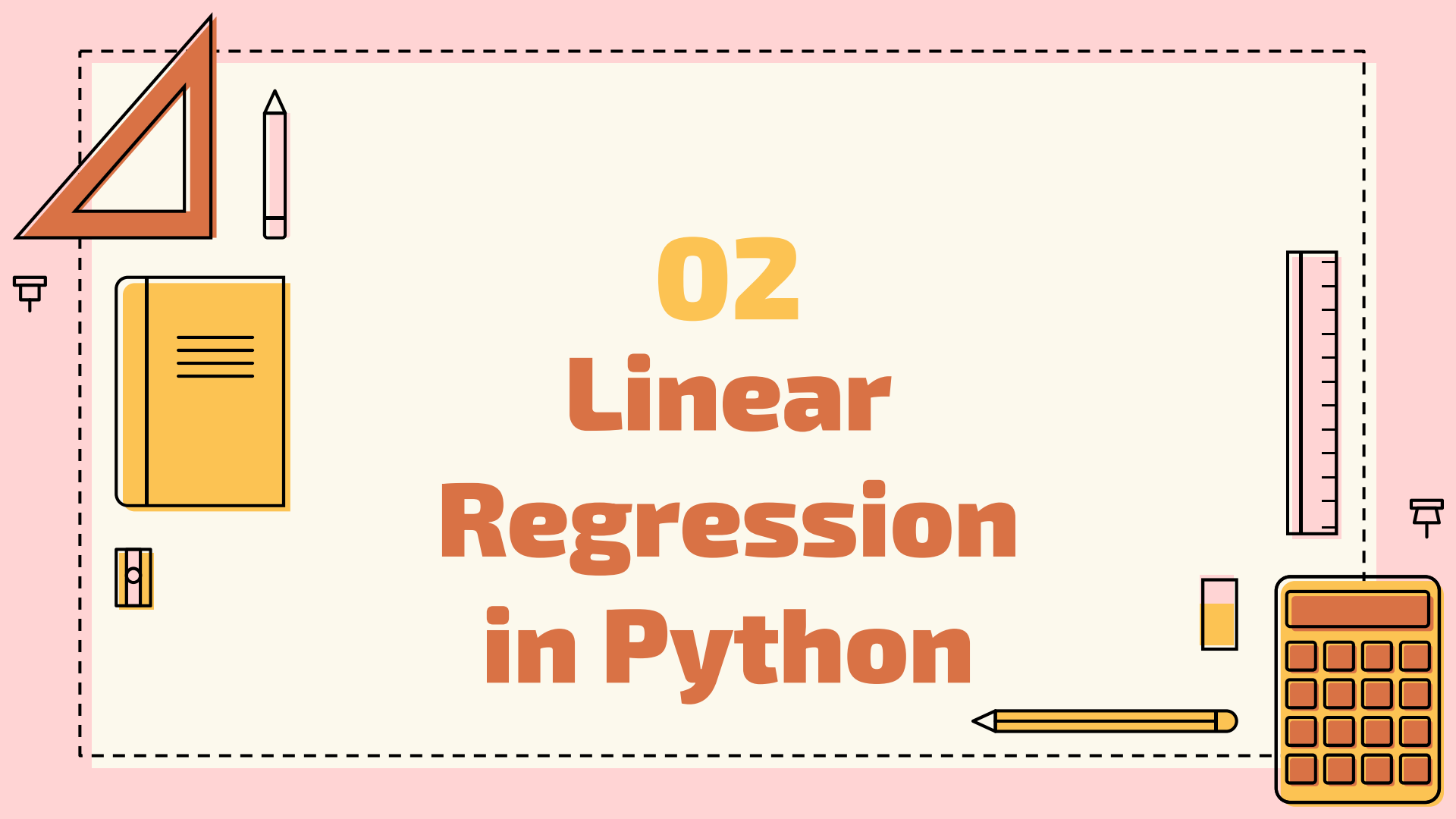
GVLMA Library of Assumptions

	Value	p-value	Decision
Global Stat	16.8650	2.053e-03	*Assumptions NOT satisfied!*
Skewness	0.2130	6.444e-01	*Assumptions acceptable.*
Kurtosis	0.1223	7.266e-01	*Assumptions acceptable.*
Link Function	15.4224	8.596e-05	*Assumptions NOT satisfied!*
Heteroscedasticity	1.1073	2.927e-01	*Assumptions acceptable.*

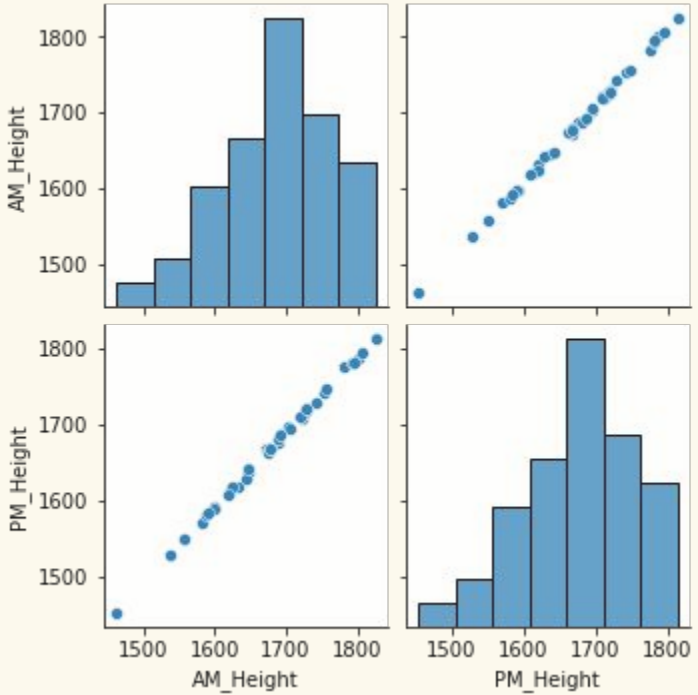
Testing for Outliers



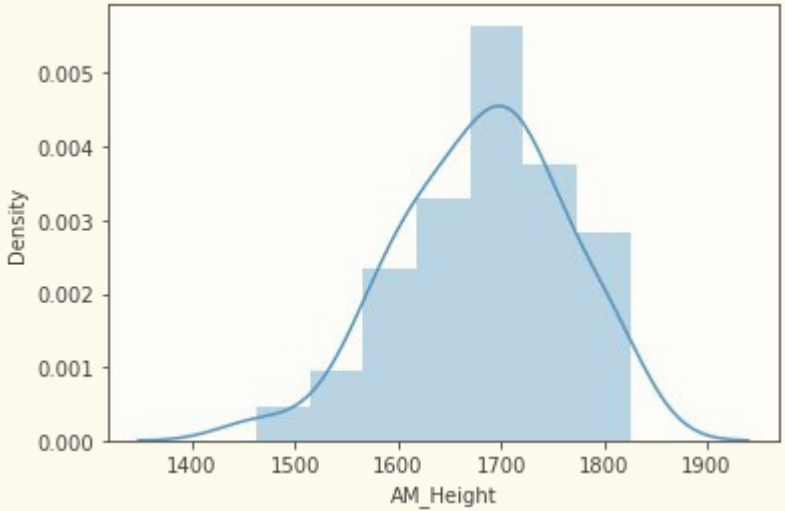
02
Linear Regression in Python

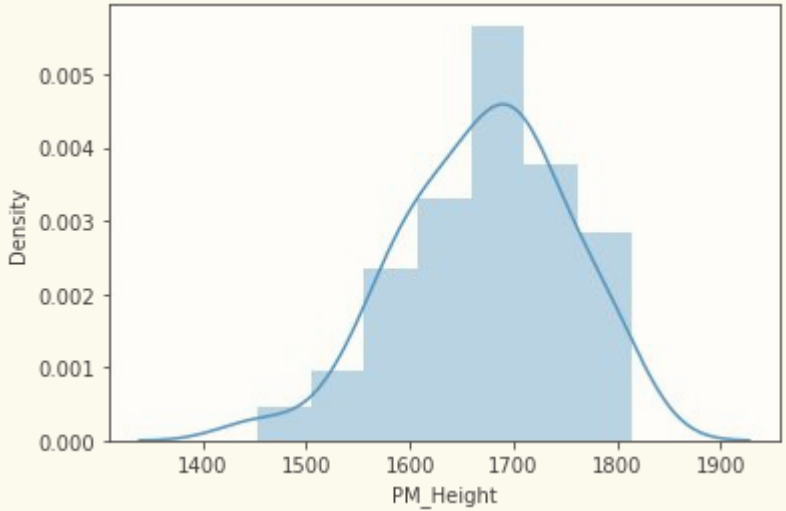


Plotting for Linearity

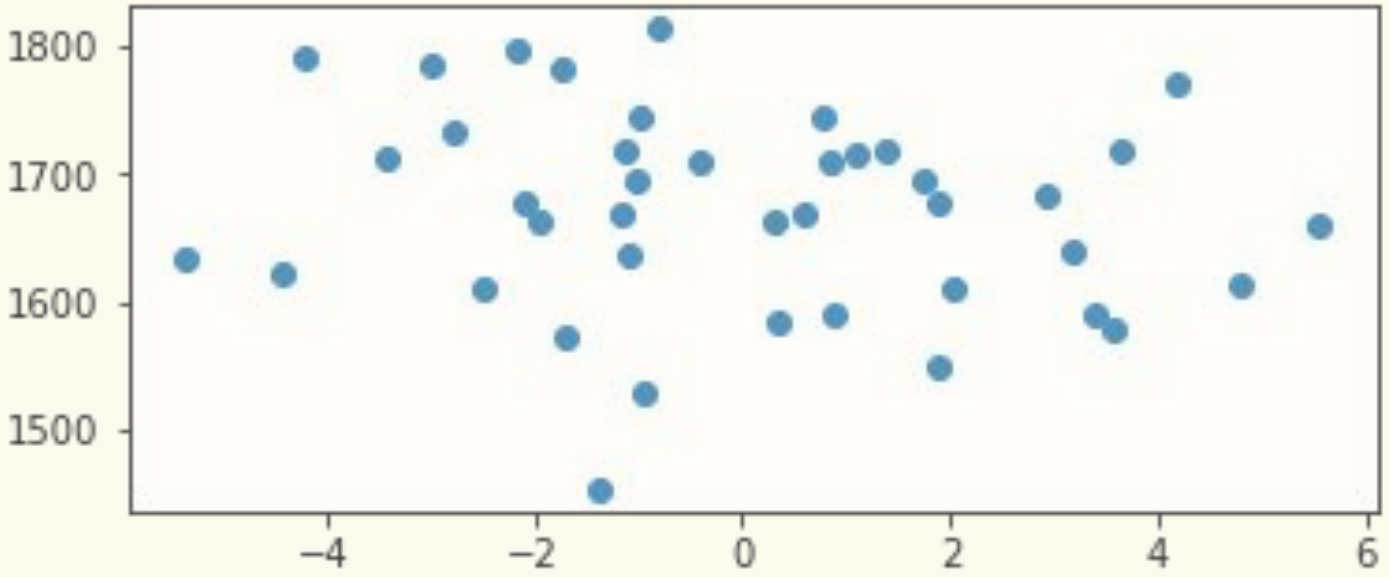


Plotting with Normal Distribution Lines





Graphing the residuals



Regression Results Summary

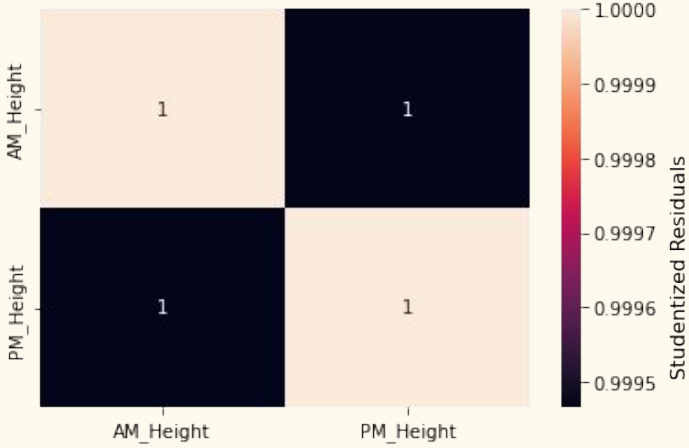
OLS Regression Results

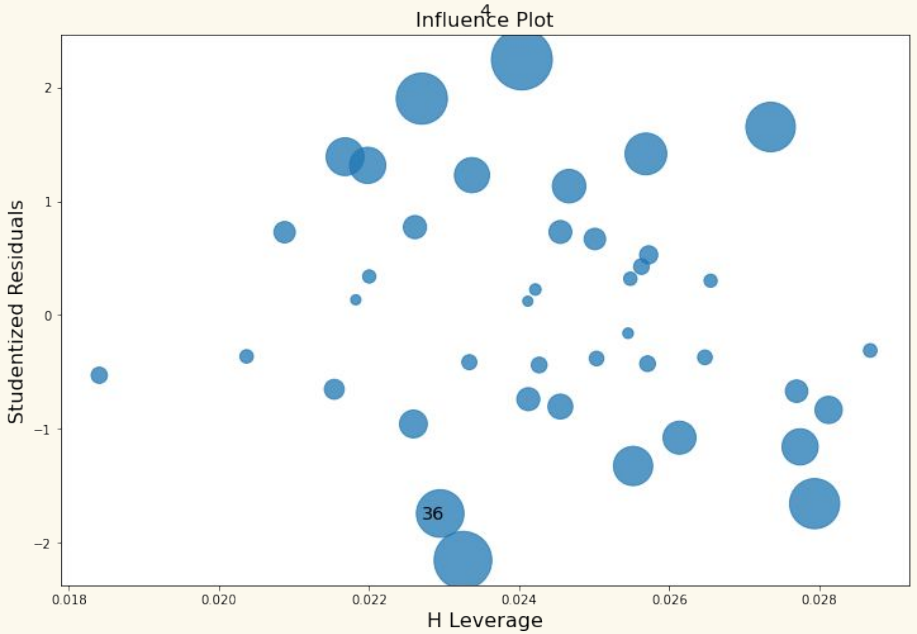
Dep. Variable:	PM_Height	**R-squared (uncentered):**	0.983
Model:	OLS	**Adj. R-squared (uncentered):**	0.983
Method:	Least Squares	**F-statistic:**	2347.
Date:	Thu, 27 Oct 2022	**Prob (F-statistic):**	3.84e-37
Time:	00:37:36	**Log-Likelihood:**	-278.68
No. Observations:	41	**AIC:**	559.4
Df Residuals:	40	**BIC:**	561.1
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
x1	1.8e-09	3.71e-11	48.450	0.000	1.72e-09	1.87e-09

Omnibus:	0.340	**Durbin-Watson:**	1.555
Prob(Omnibus):	0.844	**Jarque-Bera (JB):**	0.516
Skew:	-0.105	**Prob(JB):**	0.773
Kurtosis:	2.492	**Cond. No.**	1.00

Looking for Outliers





03
Conclusions

Conclusions

- The data looks linear when plotted
- The data seems heteroscedastic...
- But has homoscedasticity when tested
- There are several outliers

- If not working in strict statistical analysis, my preferred coding language is Python!

Thanks!

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**