GROUP NUMBER	1
TEAM MEMBERS	ADESOLA ABIMBOLA DARPAN CHAUDHARY
TITLE	Project Report

README

- Run python FinalProject.py in terminal

(a) Create a line plot of the Value for each month of your data set from 1989-2002 and 2009-2018. Include the plot in your report. Note: There is a gap in your data between 2002 and 2009.

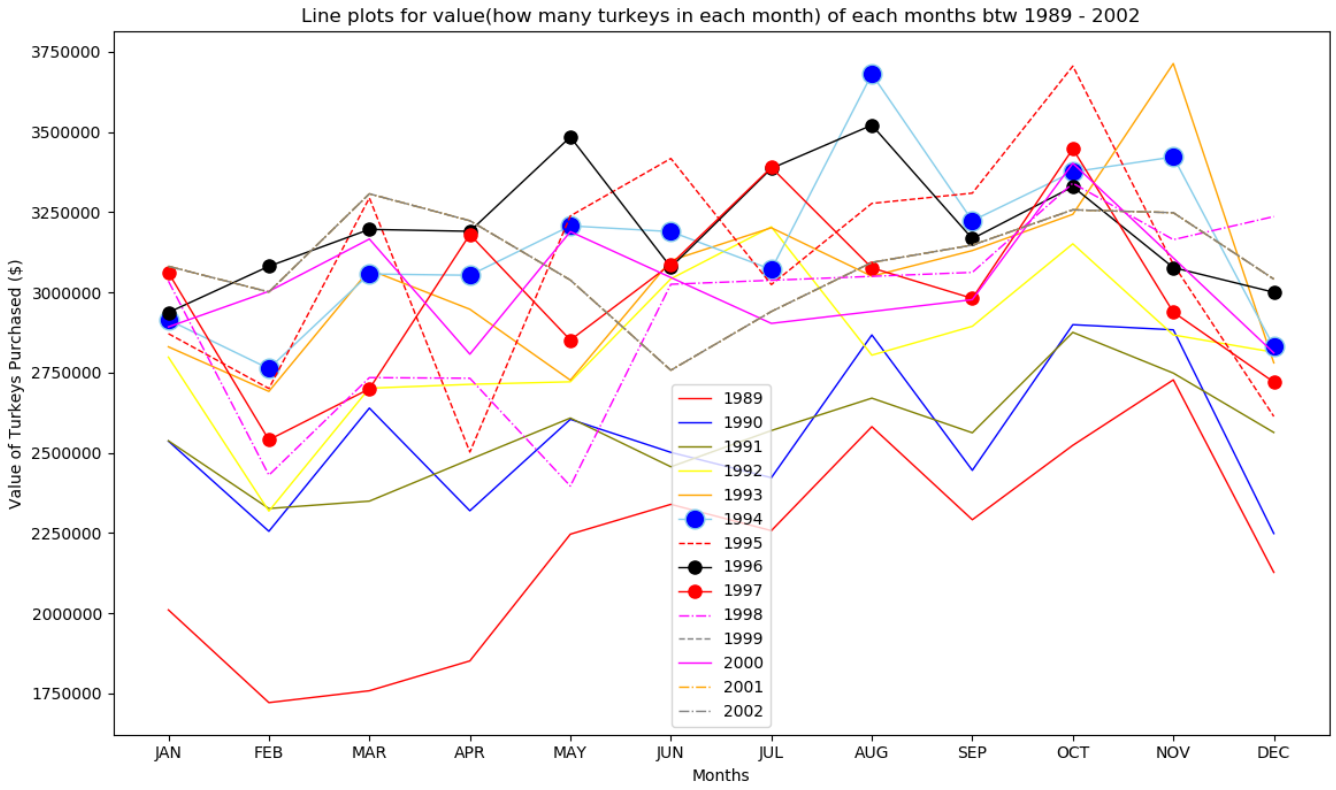


Figure 1

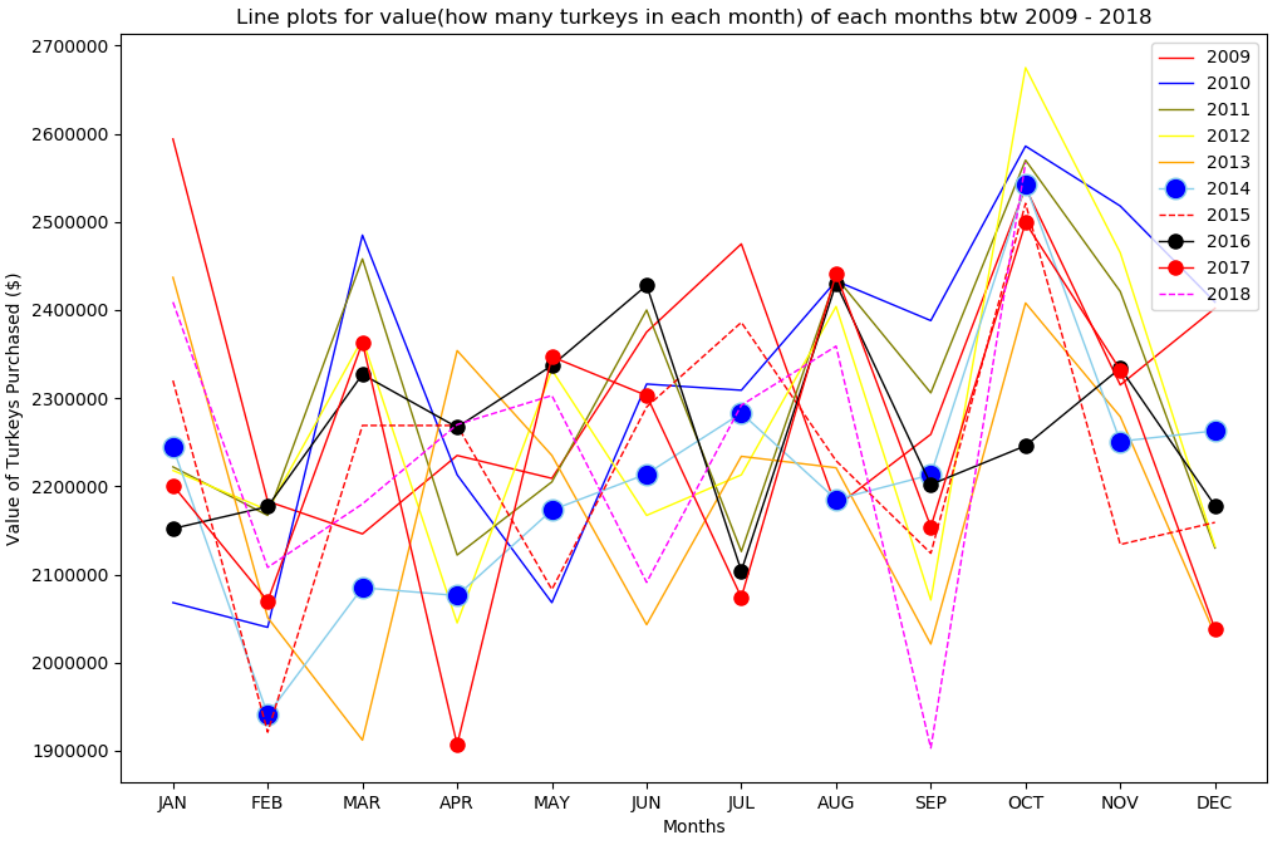


Figure 2

From Fig. 1 and Fig. 2, we can see that during the months of September and October, there is a pattern of increased value of turkeys purchased, which makes sense because there is usually a rise in turkey demand around the month of November which is the month of the thanksgiving celebration

(c) Report mean and median of the Value grouped by year

1989		1997		2011	
Mean	Median	Mean	Median	Mean	Median
2202583.33	2251500	2997416.67	3021000	2297000.00	2264000
1990		**1998**		**2012**	
Mean	Median	Mean	Median	Mean	Median
2551500.00	2518500	2915400.00	3030000	2271833.33	2215500
1991		**1999**		**2013**	
Mean	Median	Mean	Median	Mean	Median
2561833.33	2562500	3094416.67	3087000	2185583.33	2227500
1992		**2000**		**2014**	
Mean	Median	Mean	Median	Mean	Median
2835500.00	2808500	3025700.00	2990000	2205833.33	2213500
1993		**2001**		**2015**	
Mean	Median	Mean	Median	Mean	Median
3040000.00	3060000	3089250.00	3072500	2225416.67	2249000
1994		**2002**		**2016**	
Mean	Median	Mean	Median	Mean	Median
3148583.33	3130000	2866111.11	2855000	2265166.67	2256500
1995		**2009**		**2017**	
Mean	Median	Mean	Median	Mean	Median
3086416.67	3162500	2325750.00	2287000	2227416.67	2252000
1996		**2010**		**2018**	
Mean	Median	Mean	Median	Mean	Median
3203833.33	3178500	2319416.67	2352000	2248200.00	2281000

Part 3

(a) For just the data from 2017, fit a linear regression to your data for the months January – October

Here we encoded the months as integer value with jan =1 upto dec = 12 for 2017 record. We used the data from jan to oct for fitting the linear regression.

value of beta is: 9807.69230769
value of alpha is: 2163666.66667

Y_hat = x*beta + alpha

(b) Using your linear fit, predict the value of turkeys as described for November

Predicted value for november is: 2271551.28205

(c) Compute the absolute error between your predicted value and the actual value of turkeys slaughtered in Virginia in Nov 2017

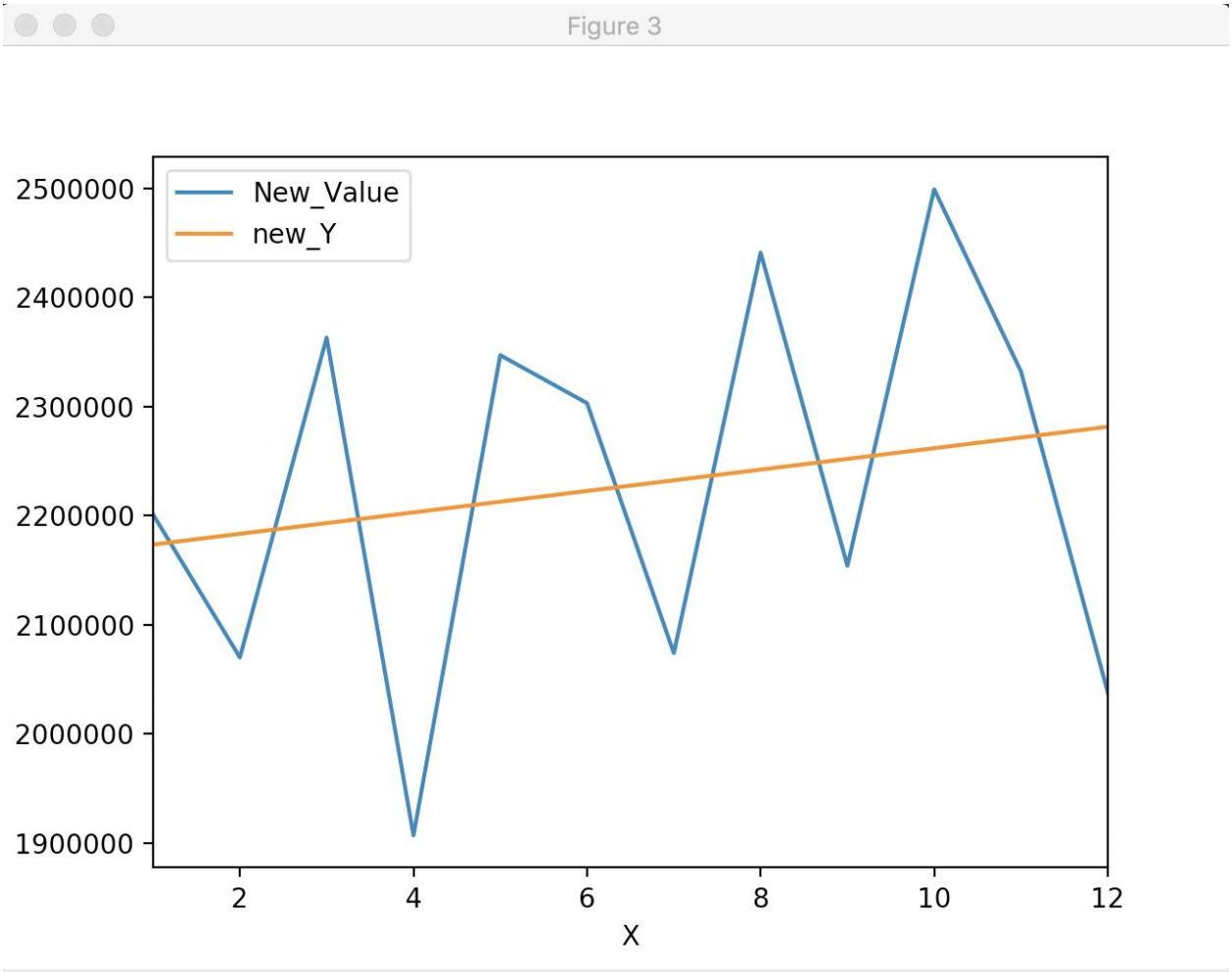
Absolute error between your predicted value and the actual value of turkeys slaughtered in Virginia in Nov 2017 is: 60448.7179487

The predicted value has a 2.66% error compared to the actual value.

(d) Compute the coefficient of determination, or R^2 value, to determine how well your model fits your data.

Coefficient of Determination is: 0.0947031374907
R-squared value is 9.4% which means that the linear regression cannot fit the graph properly

(e) Plot a line plot of Values from 2017 along with the linear fit.



As we can see in the graph there is high variance in the value of each month and linear regression cannot fit the data properly hence explaining the R squared value.