TidyTuesday Week 39

Pranay Gundam

Friday 4th October, 2024

Contents

1 Weekly Summary

The most interesting regression this week seems like the 09/25 day. Not as much because of the regression itself, but more so because of the relationship between the data that we can see on the graph itself. Specifically the two series are Nonfarm "Private Goods - Producing Large Payroll Employment" and "Mortgage Debt Outstanding by Type of Holder: Federal and Related Agencies". There seems to be a clear discontinuity with the omrtgage Dept series, where dept is only either above 5 million or below 1 million and there is no outstanding mortgage dept above 5 million on the same years that the private goods series is above 7500. I'm not so sure what else there is to say about this without more investigation but it would be interesting to see if this relationship was brought about by some nuance in the years in which this data for either series occured or if there really is some roundabout story to delve into.

2 Date: 2024-09-23

Series ID: ACOILWTICO

This series is titled Crude Oil Prices: West Texas Intermediate (WTI) - Cushing, Oklahoma and has a frequency of Annual. The units are Dollars per Barrel and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1986-01-01 and the observation end date is 2023-01-01.The popularity of this series is 22.

Series ID: STTMINWGNY

This series is titled State Minimum Wage Rate for New York and has a frequency of Annual. The units are Dollars per Hour and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1968-01-01 and the observation end date is 2024-01-01.The popularity of this series is 37.

2.1 Regression Tables and Plots

Dep. Variable:	value_fred_STTMINWGNY	R-squared:	0.444
Model:	OLS	Adj. R-squared:	0.428
Method:	Least Squares	F-statistic:	28.74
Date:	Mon, 23 Sep 2024	Prob (F-statistic):	4.97e-06
Time:	13:31:51	Log-Likelihood:	-84.492
No. Observations:	38	AIC:	173.0
Df Residuals:	36	BIC:	176.3
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P> |t|	[0.025	0.975]
const	3.3324	0.716	4.652	0.000	1.880	4.785
value_fred_ACOILWTICO	0.0699	0.013	5.361	0.000	0.043	0.096

Omnibus:	8.225	**Durbin-Watson:**	0.249
Prob(Omnibus):	0.016	**Jarque-Bera (JB):**	7.567
Skew:	1.082	**Prob(JB):**	0.0227
Kurtosis:	3.309	**Cond. No.**	106.

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
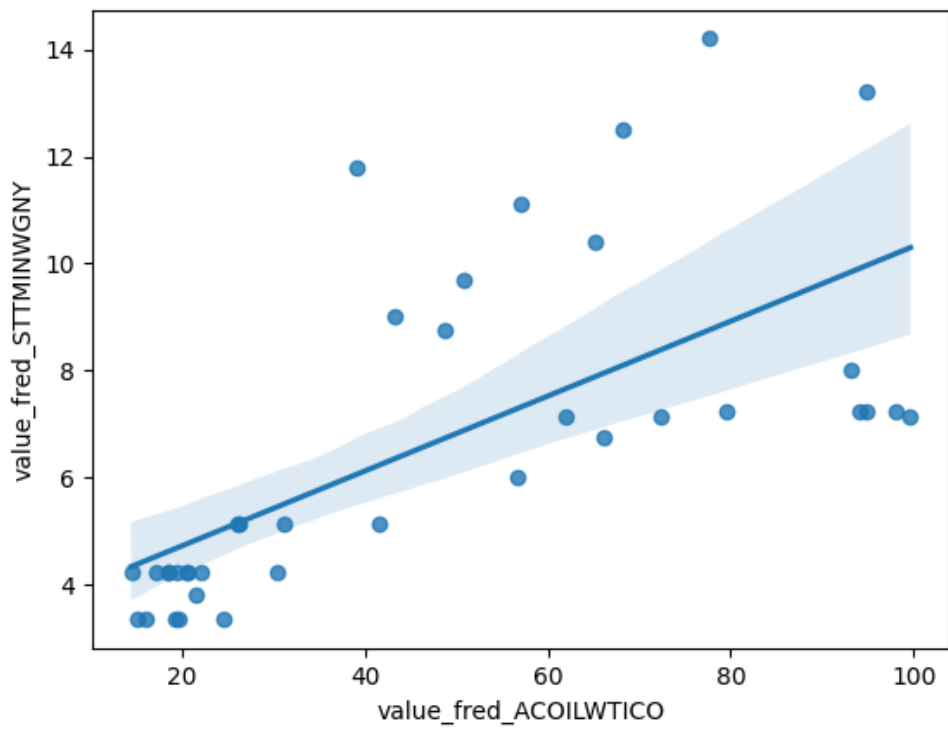
Figure 1: Regression Plot for 2024-09-23

3 Date: 2024-09-24

Series ID: CEU9000000010

This series is titled Women Employees, Government and has a frequency of Monthly. The units are Thousands of Persons and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1964-01-01 and the observation end date is 2024-08-01.The popularity of this series is 1.

Series ID: DLTRUCKSNSA

This series is titled Motor Vehicle Retail Sales: Domestic Light Weight Trucks and has a frequency of Monthly. The units are Thousands of Units and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1967-01-01 and the observation end date is 2024-08-01.The popularity of this series is 4.

3.1 Regression Tables and Plots

Dep. Variable:	value_fred_DLTRUCKSNSA	R-squared:	0.747
Model:	OLS	Adj. R-squared:	0.747
Method:	Least Squares	F-statistic:	2038.
Date:	Tue, 24 Sep 2024	Prob (F-statistic):	3.71e-208
Time:	22:43:06	Log-Likelihood:	-4259.9
No. Observations:	692	AIC:	8524.
Df Residuals:	690	BIC:	8533.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P> |t|	[0.025	0.975]
const	-297.4786	16.845	-17.660	0.000	-330.552	-264.406
value_fred_CEU9000000010	0.0728	0.002	45.147	0.000	0.070	0.076

Omnibus:	20.866	Durbin-Watson:	0.443
Prob(Omnibus):	0.000	Jarque-Bera (JB):	33.310
Skew:	-0.239	Prob(JB):	5.85e-08
Kurtosis:	3.963	Cond. No.	4.05e+04

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 4.05e+04. This might indicate that there are strong multicollinearity or other numerical problems.
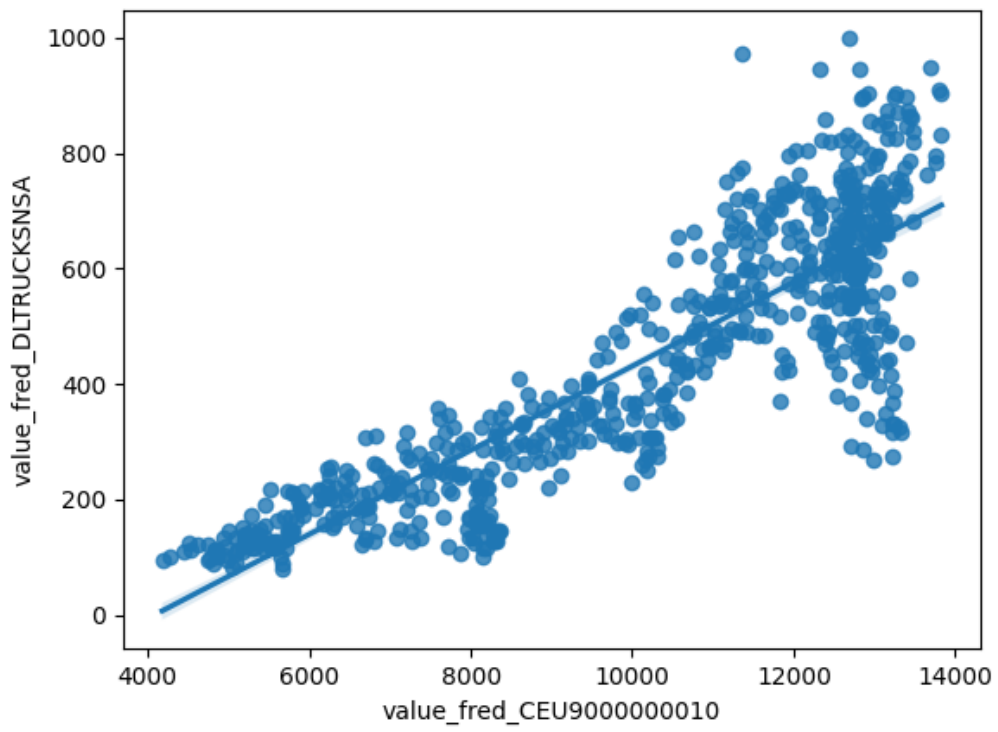
Figure 2: Regression Plot for 2024-09-24

4 Date: 2024-09-25

Series ID: NPPGPL2

This series is titled Nonfarm Private Goods - Producing Large Payroll Employment (1000+) (DISCONTINUED) and has a frequency of Monthly. The units are Thousands and the seasonal adjustment is Seasonally Adjusted.The observation start date is 2005-01-01 and the observation end date is 2022-05-01.The popularity of this series is 1.

Series ID: MDOTHFRA

This series is titled Mortgage Debt Outstanding by Type of Holder: Federal and Related Agencies (DISCONTINUED) and has a frequency of Quarterly, End of Period. The units are Millions of Dollars and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1949-01-01 and the observation end date is 2019-07-01.The popularity of this series is 3.

4.1 Regression Tables and Plots

Dep. Variable:	value_fred_MDOTHFRA	**R-squared:**	0.350
Model:	OLS	**Adj. R-squared:**	0.339
Method:	Least Squares	**F-statistic:**	30.71
Date:	Wed, 25 Sep 2024	**Prob (F-statistic):**	7.98e-07
Time:	10:06:53	**Log-Likelihood:**	-929.62
No. Observations:	59	**AIC:**	1863.
Df Residuals:	57	**BIC:**	1867.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P> |t|	[0.025	0.975]
const	2.218e+07	3.35e+06	6.616	0.000	1.55e+07	2.89e+07
value_fred_NPPGPL2	-2599.1459	468.991	-5.542	0.000	-3538.285	-1660.007

Omnibus:	6.887	**Durbin-Watson:**	0.116
Prob(Omnibus):	0.032	**Jarque-Bera (JB):**	6.256
Skew:	-0.782	**Prob(JB):**	0.0438
Kurtosis:	3.312	**Cond. No.**	1.07e+05

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 1.07e+05. This might indicate that there are strong multicollinearity or other numerical problems.
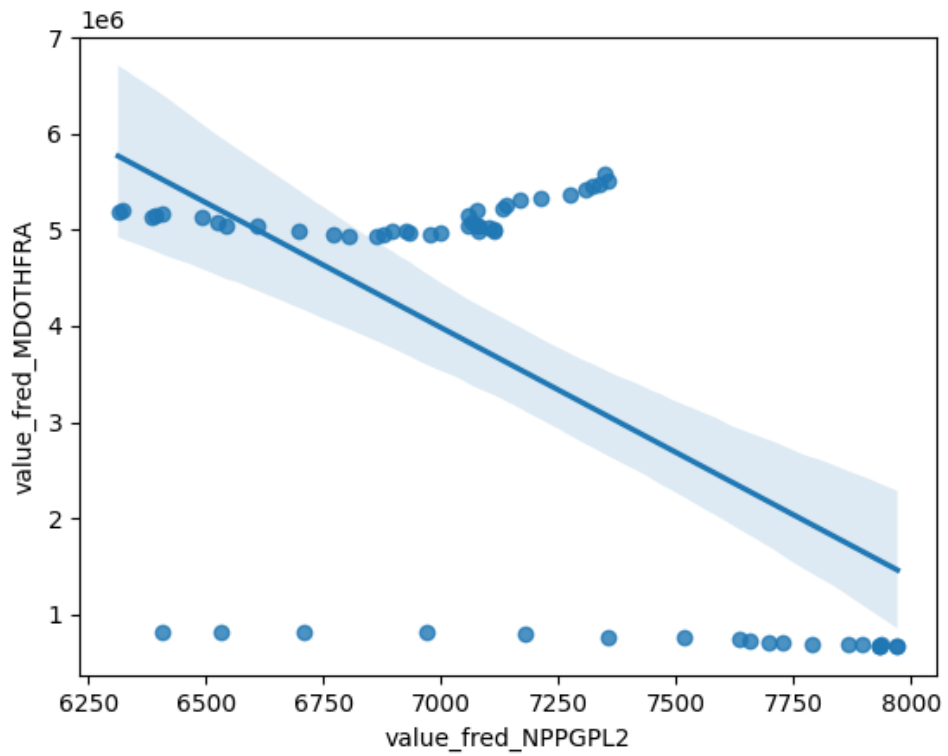
Figure 3: Regression Plot for 2024-09-25

5   Date: 2024-09-26

Series ID: MESTFININSRGSP

This series is titled Real Gross Domestic Product: Finance and Insurance (52) in the Mideast BEA Region and has a frequency of Annual. The units are Millions of Chained 2017 Dollars and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1997-01-01 and the observation end date is 2023-01-01.The popularity of this series is 1.

Series ID: W514RC1A027NBEA

This series is titled Government current receipts: Excluding imputations and has a frequency of Annual. The units are Billions of Dollars and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1929-01-01 and the observation end date is 2022-01-01.The popularity of this series is 1.

5.1   Regression Tables and Plots

Dep. Variable:	value_fred_W514RC1A027NBEA	R-squared:	0.606
Model:	OLS	Adj. R-squared:	0.589
Method:	Least Squares	F-statistic:	36.87
Date:	Thu, 26 Sep 2024	Prob (F-statistic):	2.86e-06
Time:	12:14:24	Log-Likelihood:	-211.35
No. Observations:	26	AIC:	426.7
Df Residuals:	24	BIC:	429.2
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P> |t|	[0.025	0.975]
const	-3061.1692	1234.241	-2.480	0.021	-5608.518	-513.820
value_fred_MESTFININSRGSP	0.0183	0.003	6.072	0.000	0.012	0.024

Omnibus:	3.731	Durbin-Watson:	0.868
Prob(Omnibus):	0.155	Jarque-Bera (JB):	2.374
Skew:	0.721	Prob(JB):	0.305
Kurtosis:	3.332	Cond. No.	3.02e+06

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 3.02e+06. This might indicate that there are strong multicollinearity or other numerical problems.
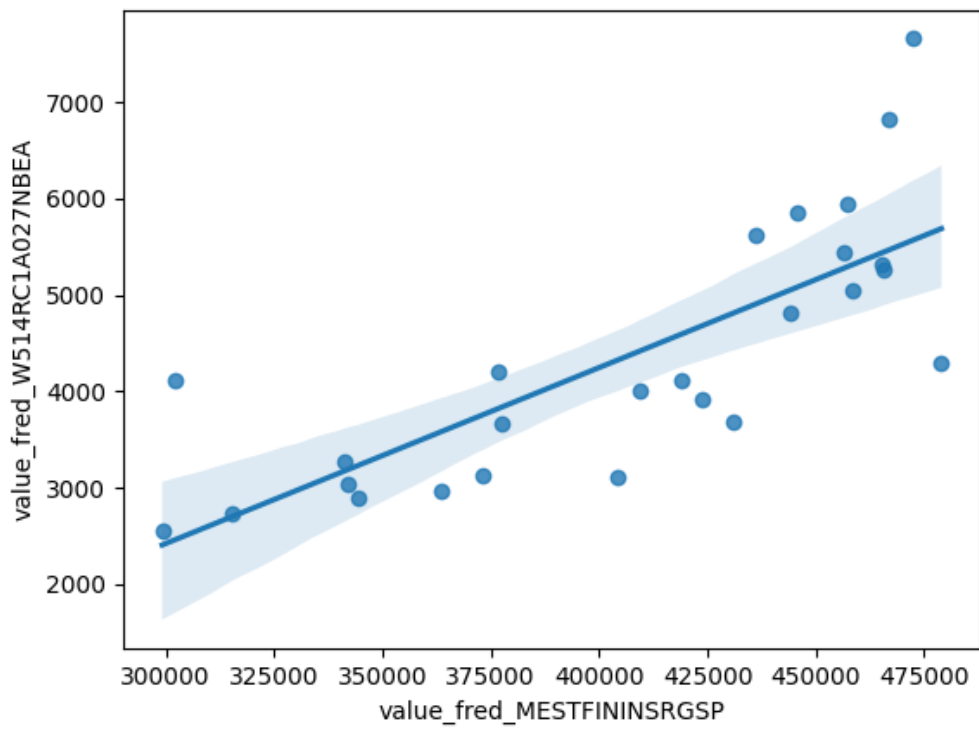
Figure 4: Regression Plot for 2024-09-26

6   Date: 2024-09-28

Series ID: PECILBU18MA25023A647NCEN

This series is titled 90

Series ID: DLTRUCKSNSA

This series is titled Motor Vehicle Retail Sales: Domestic Light Weight Trucks and has a frequency of Monthly. The units are Thousands of Units and the seasonal adjustment is Not Seasonally Adjusted.The observation start date is 1967-01-01 and the observation end date is 2024-08-01.The popularity of this series is 4.

6.1   Regression Tables and Plots

Dep. Variable:	value_fred_DLTRUCKSNSA	R-squared:	0.354
Model:	OLS	Adj. R-squared:	0.330
Method:	Least Squares	F-statistic:	14.80
Date:	Sat, 28 Sep 2024	Prob (F-statistic):	0.000663
Time:	16:52:48	Log-Likelihood:	-169.80
No. Observations:	29	AIC:	343.6
Df Residuals:	27	BIC:	346.3
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	752.4154	73.373	10.255	0.000	601.866	902.965
value_fred_PECILBU18MA25023A647NCEN	-0.0299	0.008	-3.847	0.001	-0.046	-0.014

Omnibus:	4.911	Durbin-Watson:	0.707
Prob(Omnibus):	0.086	Jarque-Bera (JB):	3.382
Skew:	-0.798	Prob(JB):	0.184
Kurtosis:	3.502	Cond. No.	4.26e+04

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 4.26e+04. This might indicate that there are strong multicollinearity or other numerical problems.
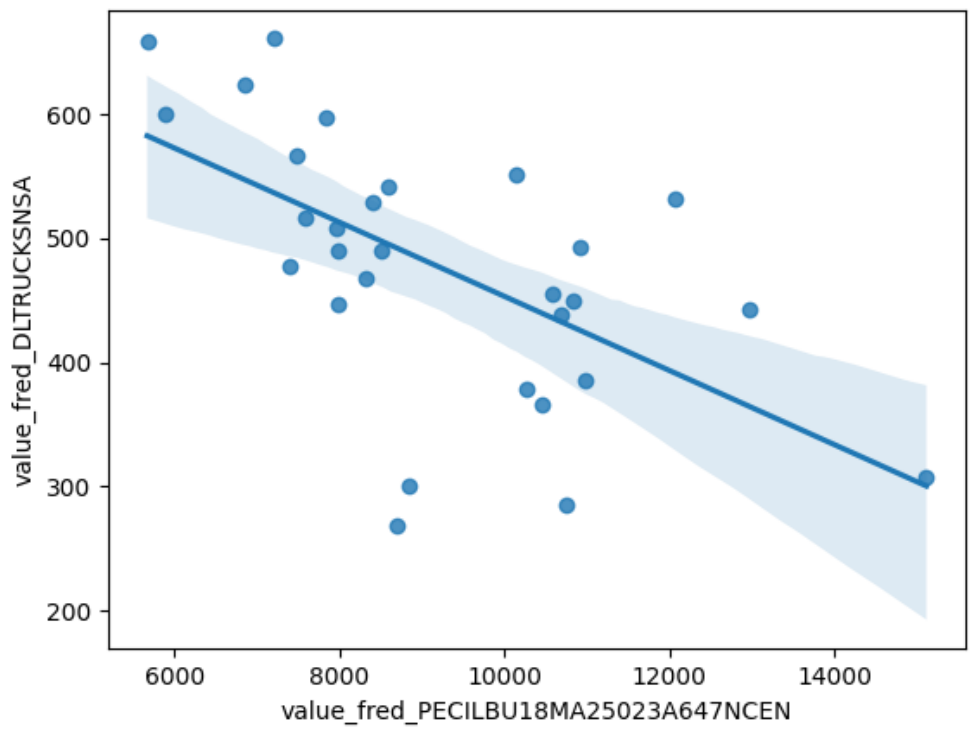
Figure 5: Regression Plot for 2024-09-28