Call Forecasting for Inbound Call Center, Sep. 2016

loadData

In [2]:

```
pwd
```

Out[2]:

```
u'c:\\users\\crucker\\desktop\\python'
```

In [19]:

```
import pandas as pd
xls_file = pd.ExcelFile('call.xls')
xls_file
```

Out[19]:

```
<pandas.io.excel.ExcelFile at 0x7ecf160>
```

In [20]:

```
xls_file.sheet_names
```

Out[20]:

```
[u'call']
```

In [21]:

```
df = xls_file.parse('call')
df.head()
```

Out[21]:

	date	calls handled
0	1	93063
1	2	83982
2	3	81342
3	4	82143
4	5	77217

cleanData

In [22]:

```
df_clean = df.dropna()
df_clean.head(1)
df_clean.ix[:,1:2].head(2)
df_clean[['date', 'calls handled']].corr().ix[:,1:2].head(2)
```

Out[22]:

	calls handled
date	-0.942662
calls handled	1.000000

buildModel

In [23]:

```
import statsmodels.api as sm
X = df_clean[['date']]
y = df_clean[['calls handled']]
X1 = sm.add_constant(X)
est = sm.OLS(y, X1).fit()
est.summary()
```

Out[23]:

OLS Regression Results

Dep. Variable:	calls handled	**R-squared:**	0.889
Model:	OLS	**Adj. R-squared:**	0.883
Method:	Least Squares	**F-statistic:**	167.5
Date:	Fri, 23 Sep 2016	**Prob (F-statistic):**	1.78e-11
Time:	09:00:40	**Log-Likelihood:**	-223.98
No. Observations:	23	**AIC:**	452.0
Df Residuals:	21	**BIC:**	454.2
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[95.0% Conf. Int.]**
const	8.861e+04	1850.243	47.891	0.000	8.48e+04 9.25e+04
date	-1746.6028	134.943	-12.943	0.000	-2027.232 -1465.974

Omnibus:	1.017	**Durbin-Watson:**	1.501
Prob(Omnibus):	0.602	**Jarque-Bera (JB):**	0.873
Skew:	-0.226	**Prob(JB):**	0.646
Kurtosis:	2.159	**Cond. No.**	28.5

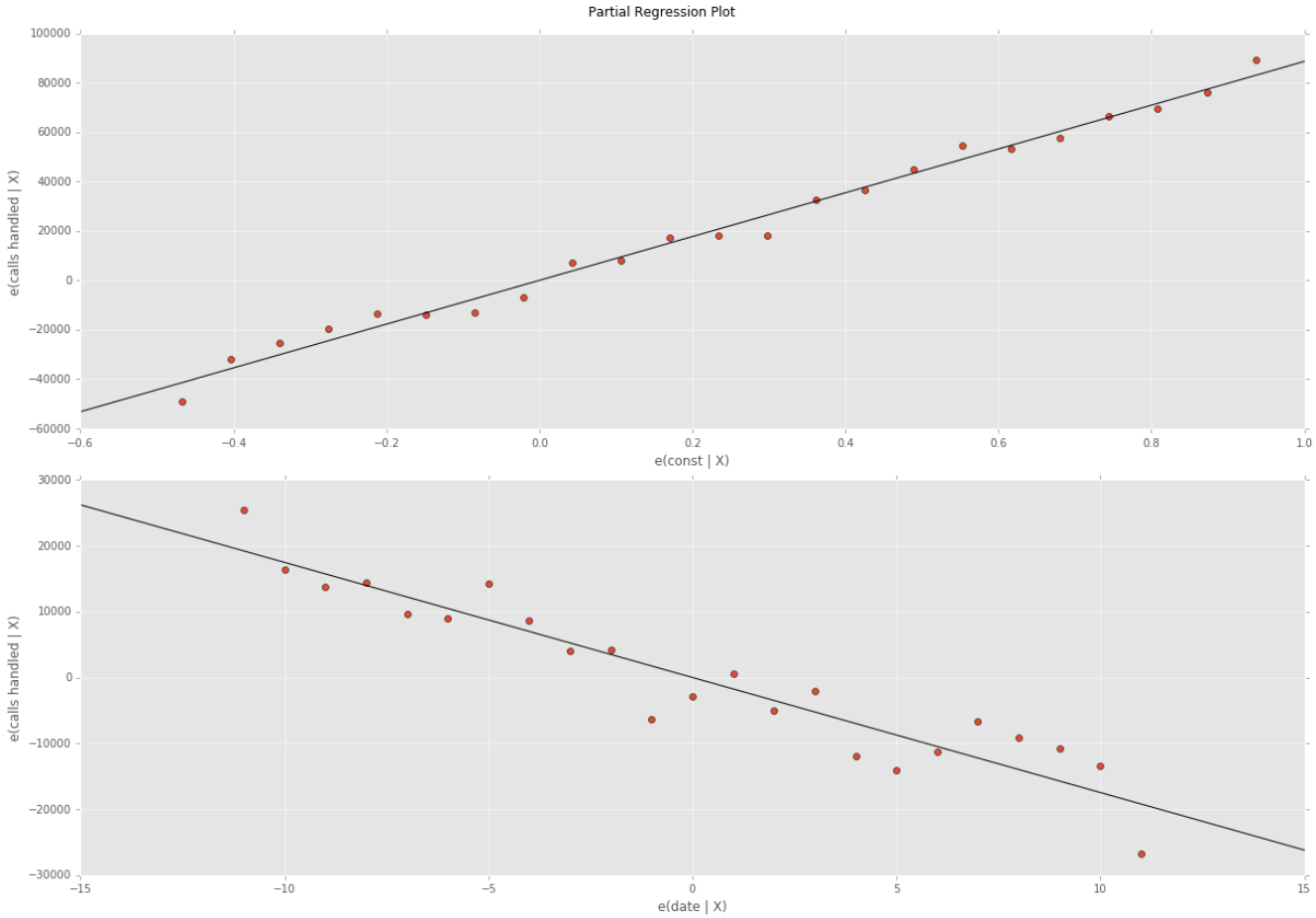
showRegressors

In [24]:

```
import matplotlib.pyplot as plt
import statsmodels.api as sm
%matplotlib inline

with plt.style.context('ggplot'):
    fig = plt.figure(figsize=(17,12))
    fig = sm.graphics.plot_partregress_grid(est, fig=fig)

plt.show()
```



writeUp

This is an Ordinary Least Squares (OLS) regression model that results in predicted values close to the observed data.

The R-squared value in the OLS Regression Results is a relative measure of fit, and improvement in the regression model results in proportional increases in R-squared as predictors are added to the regression model.

R-squared value for this model is 0.889 or 88.9%, with an adjusted R-squared value of 0.883 or 88.3%. The reliability of this regression model is reinforced by the extreme correlation coefficient $R \approx -0.94$. For a sample of size 23, a correlation of $-1.0 \leq R \leq -0.5$ is considered to be a strong negative correlation.

The data clearly suggests that the number of inbound calls is strongly dependent on time and decreasing at a steady pace of around 1,746 calls per month.

journalCitation

- Vinje, Peter (2009) "Call Forecasting for Inbound Call Center," Undergraduate Journal of Mathematical Modeling: One + Two: Vol. 2: Iss. 1, Article 7.
- DOI: http://dx.doi.org/10.5038/2326-3652.2.1.7
- Available at: http://scholarcommons.usf.edu/ujmm/vol2/iss1/7