

In [2]:

```
from datacl import Clean
import statsmodels.api as sm
import statsmodels.formula.api as smf

clean=Clean()
df=clean.load_excel("D:/Github/BigData-project/weather.xlsx")
y=df.iloc[:, :1]
x=df.iloc[:, 1:]
```

In [3]:

```
model=smf.ols("price ~ temp", data=df)
res=model.fit()
print(res.summary())
```

OLS Regression Results

```
=====
Dep. Variable:          price    R-squared:            0.031
Model:                  OLS      Adj. R-squared:       0.028
Method:                 Least Squares    F-statistic:      10.23
Date:                  Sun, 26 Aug 2018    Prob (F-statistic):  0.00152
Time:                  11:52:25    Log-Likelihood:     -2994.7
No. Observations:      326    AIC:                5993.
Df Residuals:          324    BIC:                6001.
Df Model:              1
Covariance Type:       nonrobust
=====
```

	coef	std err	t	P> t	[0.025	0.975]
Intercept	1.237e+04	210.952	58.649	0.000	1.2e+04	1.28e+04
temp	38.7403	12.110	3.199	0.002	14.916	62.564

```
=====
Omnibus:                16.455    Durbin-Watson:        1.780
Prob(Omnibus):          0.000    Jarque-Bera (JB):      24.844
Skew:                  -0.358    Prob(JB):              4.03e-06
Kurtosis:               4.147    Cond. No.:             28.1
=====
```

Warnings:

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

In [4]:

```
fig = sm.graphics.plot_regress_exog(res, "temp")
```



