Predicting Leasing Activity with Economic Indicators for Class A office markets

Apple SeeRs: Cooper Ruffing, Liane Ma, Amy Xu, Hyunjin Lee, Phillip Lin





EDA of Unemployment, Leases, and FRED data of interest



Fitting A Linear Model

term	estimate	std.error	statistic	p.value
(Intercept)	-56.883	14.217	-4.001	0.000
month	-0.103	0.088	-1.178	0.239
CBD_suburbanSuburban	-1.838	0.855	-2.151	0.032
UMCSENT	0.088	0.035	2.479	0.013
marketAustin	1.090	2.090	0.521	0.602
marketBaltimore	1.340	2.373	0.565	0.572
marketBoston	-0.568	1.647	-0.345	0.730
marketCharlotte	4.645	2.699	1.721	0.085
marketChicago Suburbs	0.313	2.147	0.146	0.884
marketDetroit	3.073	2.985	1.029	0.303
marketHouston	-0.469	1.685	-0.278	0.781
marketLos Angeles	-1.198	1.639	-0.731	0.465
marketManhattan	-1.869	1.476	-1.266	0.205
marketNashville	-1.015	3.327	-0.305	0.760
marketNorthern New Jersey	0.121	1.696	0.071	0.943
marketNorthern Virginia	0.023	1.736	0.013	0.990
marketPhiladelphia	-0.984	2.222	-0.443	0.658
marketPhoenix	0.768	2.402	0.320	0.749
marketSalt Lake City	4.309	4.127	1.044	0.297
marketSan Diego	0.483	2.006	0.241	0.810
marketSan Francisco	1.217	2.055	0.592	0.554
marketSeattle	-0.469	2.022	-0.232	0.817
marketSouth Florida	0.708	2.405	0.294	0.769
marketTampa	0.017	2.500	0.007	0.995
REAINTRATREARAT10Y	-0.760	0.483	-1.574	0.115
тси	0.602	0.155	3.879	0.000
BABANAICSRETSAUS	0.000	0.000	3.514	0.000

- Created with only the variables that were significant (α > 0.1) in the output from the first model
- Certain variables from FRED of interest, used a Drop-in Deviance test to compare nested models and identified a new subset of variables
- Pivoted to **XGBoost model** to better account for the nonlinearity and complex interactions between variables that our linear model missed

<pre>sigfit5 <- lm(lease_ch_pct ~ month + CBD_suburban + UMCSENT + market + REAINTRATREARAT10</pre>
<pre>summary(sig_fit)\$adj.r.squared</pre>
[1] 0.003304348
<pre>summary(sigfit5)\$adj.r.squared</pre>
[1] 0.003880844
AIC(sig_fit, sigfit5)
df AIC sig_fit 5 28205.24 sigfit5 28 28226.21
BIC(sig_fit, sigfit5)
df BIC sig_fit 5 28235.76 sigfit5 28 28397.16

→ suboptimal R² and RMSE indicate linear assumptions and additive structure of the model are too simplistic for complex, nonlinear dynamics we are trying to represent

XGBoost Model





Results:

- Residuals plot clustered near zero, underfits at higher values
- **RMSE**: 0.4415; **MAE**: 0.2960, **R-squared**: 0.3959



XGBoost Feature Importance (Gain > 0.5%)