Executive Summary

We have divided our project into 3 parts where in the first part we work with zip code level data, second part we work with property level data and third part we perform an analysis as to which properties are best for investment.

Firstly, we try to figure out what characteristics are associated with the rent per square foot, and we collect data of 48 different variables on zip code level, run regression models on the median rent per square foot of each zip code, and finally picked 11 variables below that we think could affect the rent per square foot, with a high adjusted R-square of 0.472.

- Had_Birth_35to50_pct
- NeverMarriedMales_20to34_pct
- Median_Earnings_thousand
- English_Speaking_Households_pct
- MovedFromOtherStates_25YearsOver_pct
- Households_OneOrMore_Computer_pct
- Num_Costco
- Num_Walmart
- Num_Kroger
- Num_Starbucks
- Num_Museums

Then we predict the rent per square foot using the model , compare it to the actual value, and calculate the difference between them.

	Zip Code	MED_RENT_PER_SQFT	Zip_Prediction	Zip_Diff	Zip_Rank
2572	87116	0.926792	2.140326	1.213533	1
2612	89109	1.347579	2.544902	1.197323	2
462	31409	1.489480	2.469451	0.979971	3
2174	79849	1.100953	2.073549	0.972596	4
2338	83422	0.857306	1.823393	0.966087	5

Hence we initially get the top 5 zip codes above with the highest potential for raising rent.

Secondly, we used property characteristics (property's age and imparting & locrating rating) to build a regression model on property level (adj R-squared is 0.291). Using this model, we predict the rent per square foot for each property and rank each property based on the difference between the predicted value and actual value. The below graph shows the TOP 5 properties that have the biggest difference.

	PROPERTY_ID	Zip Code	PROPERTY_SQFT	RENT_PER_SQFT	Prof_Prediction	Prop_Diff	Prop_Rank
10883	61_1508	78258	480460	1.295508	2.451608	1.156099	1
5723	124_97	30809	260404	1.087664	2.196034	1.108370	2
538	14_6466	89149	112404	1.162734	2.260548	1.097814	3
4794	54_188	72211	571835	1.392445	2.431280	1.038834	4
1330	104_107	37803	53450	0.564079	1.584505	1.020426	5

On the third part, we collect data of total population and unemployment of recent years on zip code level and calculate the population growth rate and unemployment rate to reflect the demand of rent.

Population_Avg_Yearly_Growth	Population_Rank	UnemploymentRate_Avg_Yearly_Growth	Unemployment_Rank
0.008042	13814	-0.057738	13066
0.008042	13814	-0.057738	13066
0.008042	13814	-0.057738	13066

Finally, we merged the above three parts to select ideal properties for investment. We calculated the estimated dollars we can gain for each property, using the following formula.

Estimated dollars = ((Predicted rent per sqft from part 1 + Predicted rent per sqft from part 2) / 2 - Actual rent per sqft) * sqft

In order to get the ideal properties , we set restrcitions:
(1) All recommended properties have to be in Zip Code with positive population growth and negative unemployment rate growth.
(2) All recommended properties have to have a positive difference between predicted rent per square foot and actual value from part 1 and part 2.

Then we come up with two methods to get the five properties that we recommend.

(1) Randomly set the threshold of the rank of zip code in part 1 and the rank of property in part 2 until we get only 5 properties:
Zip_Rank<30, Prop_Rank<163

PROPERTY_ID	Zip Code	Prop_Rank	Zip_Rank	Dollar
118_89	39194	29	21	61476.236345
73_361	35406	102	22	197910.693556
54_140	72207	143	20	53400.381846
89_451	79927	155	7	123386.282647
119_1088	36527	162	27	147726.931134

(2) Set thresholds for Zip_Rank and Prop_Rank to be top 10%, select 5 properties with highest 'Dollar'

PROPERTY_ID	Zip Code	Prop_Rank	Zip_Rank	Dollar
178_1110800	72712	269	29	357707.714324
94_166	74075	256	35	304092.540451
46_1554	27519	185	24	284817.710109
94_87	74075	339	35	282542.017736
178_1110703	72712	232	29	261837.641540

We agree that the second one is more reasonable.

Hence, our recommendation is to invest in properties 178_1110800, 94_166, 46_1554, 94_87 and 178_1110703 which have the highest potential for return of investment.

PROPERTY_ID	MARKET_NAME	COUNTY_NAME	PROPERTY_NAME	PROPERTY_CITY	PROPERTY_STATE	PROPERTY_UNITS	PROPERTY_SQFT	PROPERTY_IMPRATING	PROPERTY_LOCRATING
178_1110800	NW Arkansas	Benton	Walton Crossing	Bentonville	AR	600	520200	A-	B-
94_166	Oklahoma City	Payne	Stillwater Student Living	Stillwater	OK	231	439359	B	B-
46_1554	Raleigh - Durham	Wake	Bradford, The	Cary	NC	390	437947	A+	A-
94_87	Oklahoma City	Payne	Links at Stillwater I, The	Stillwater	OK	492	423128	B	B
178_1110703	NW Arkansas	Benton	Links at Bentonville, The	Bentonville	AR	432	358576	A-	B