

# Boston Real Estate Rents Analysis

## Team 4

Wenjin Cao  
Jianxing Lu  
Qiaomin Ling

# Project Description

- In our project, we can predict rents in Boston based on different arguments such as number of bedrooms/bathrooms, building type, location (zip code, public transportation), etc.

We also did a vivid and interactive map-based visualization and analysis of Boston rents in accordance with specific neighborhoods and dates.

# 3 Kinds of Dataset Sources

- (1) Zillow Data

<http://www.zillow.com/research/data/#rental-data>

We used the monthly rental data of different neighborhoods in Boston from 2013/01 to 2016/06.

Parameters: Median rent (per neighborhood)

RegionName	X2015.01	X2015.02	X2015.03	X2015.04	X2015.05	X2015.06	X2015.07	X2015.08	X2015.09	X2015.10	X2015.11	X2015.12
Jamaica Plain	1325.0	1337.5	1197.5	1275.0	1262.5	1275	1300.0	1400	1425.0	1475.0	1625.0	
Brighton	1375.0	1397.5	1395.0	1395.0	1400.0	1395	1450.0	1425	1425.0	1400.0	1395.0	
South Boston	2250.0	2400.0	2299.5	2321.0	2348.0	2325	2405.5	2400	2375.0	2375.0	2302.5	
South End	1862.5	1950.0	1750.0	1750.0	1725.0	1675	1750.0	1800	1950.0	2000.0	2000.0	
Allston	1395.0	1400.0	1395.0	1395.0	1400.0	1400	1400.0	1400	1395.0	1395.0	1395.0	
Fenway	1650.0	1625.0	1650.0	1650.0	1650.0	1650	1675.0	1655	1650.0	1625.0	1645.0	
Back Bay	1850.0	1800.0	1725.0	1750.0	1700.0	1745	1750.0	1795	1800.0	1800.0	1800.0	
Kenmore	1550.0	1550.0	1600.0	1600.0	1625.0	1625	1625.0	1610	1595.0	1575.0	1600.0	
Mission Hill	1300.0	1412.5	1450.0	1484.5	1495.0	1400	1400.0	1450	1510.0	1550.0	1400.0	
North End	1850.0	1850.0	1800.0	1775.0	1787.5	1700	1750.0	1800	1700.0	1799.0	1900.0	
Beacon Hill	1750.0	1750.0	1750.0	1750.0	1750.0	1750	1700.0	1700	1725.0	1797.5	1800.0	
West End	2000.0	2100.0	2100.0	2047.5	2100.0	2050	2200.0	2240	2267.5	2275.0	2327.5	

# 3 Kinds of Dataset Sources

- (2) Rental Data from Individual Third Party  
[http://www.jefftk.com/apartment\\_prices/data-listing](http://www.jefftk.com/apartment_prices/data-listing)

We used the monthly rental data of different apartments in Boston from 2013/02 to 2016/07.

Parameters: Longitude and latitude of the house, number of bedrooms, rent (per house/apt).

# 3 Kinds of Dataset Sources

- (2) Rental Data from Individual Third Party

[http://www.jefftk.com/apartment\\_prices/data-listing](http://www.jefftk.com/apartment_prices/data-listing)

```
2013-02-18.geojson
~/Downloads/adsfinal/src/main/webapp/resources/2013-02-18.geojson
1 [{"type": "FeatureCollection", "features": [{"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1569887, 42.2633768]}, "properties":
2 {"Bedroom": 1, "Price": 500}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.086491, 42.322115]}, "properties":
3 {"Bedroom": 1, "Price": 500}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1380832, 42.3565723]}, "properties":
4 {"Bedroom": 1, "Price": 550}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0729931, 42.2941909]}, "properties":
5 {"Bedroom": 4, "Price": 550}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1577728, 42.373295]}, "properties":
6 {"Bedroom": 1, "Price": 600}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1292802, 42.3526067]}, "properties":
7 {"Bedroom": 0, "Price": 650}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1034098, 42.3306177]}, "properties":
8 {"Bedroom": 1, "Price": 650}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.091694, 42.329816]}, "properties":
9 {"Bedroom": 1, "Price": 650}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0372718, 42.3306171]}, "properties":
10 {"Bedroom": 0, "Price": 675}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.15947, 42.30114]}, "properties":
11 {"Bedroom": 1, "Price": 675}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1402439, 42.404167]}, "properties":
12 {"Bedroom": 1, "Price": 675}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1438962, 42.341094]}, "properties":
13 {"Bedroom": 4, "Price": 675}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0021, 42.3908]}, "properties":
14 {"Bedroom": 3, "Price": 700}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.168916, 42.343392]}, "properties":
15 {"Bedroom": 0, "Price": 750}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1504562, 42.3549199]}, "properties":
16 {"Bedroom": 1, "Price": 750}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.168916, 42.343392]}, "properties":
17 {"Bedroom": 3, "Price": 750}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1121842, 42.3327401]}, "properties":
18 {"Bedroom": 0, "Price": 800}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0271377, 42.334904]}, "properties":
19 {"Bedroom": 0, "Price": 800}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.059712, 42.3766569]}, "properties":
20 {"Bedroom": 0, "Price": 800}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1054599, 42.3496119]}, "properties":
21 {"Bedroom": 1, "Price": 800}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1628845, 42.2768106]}, "properties":
22 {"Bedroom": 1, "Price": 800}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.082931, 42.33247]}, "properties":
23 {"Bedroom": 1, "Price": 800}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.029324, 42.3651692]}, "properties":
24 {"Bedroom": 1, "Price": 850}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0419176, 42.4339217]}, "properties":
25 {"Bedroom": 1, "Price": 850}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1692031, 42.3441413]}, "properties":
26 {"Bedroom": 2, "Price": 850}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0838132, 42.3356914]}, "properties":
27 {"Bedroom": 1, "Price": 875}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1163386, 42.3485271]}, "properties":
28 {"Bedroom": 0, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1401128, 42.3363406]}, "properties":
29 {"Bedroom": 0, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1464497, 42.2987312]}, "properties":
30 {"Bedroom": 0, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0097, 42.4104]}, "properties":
31 {"Bedroom": 1, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0103177, 42.2559937]}, "properties":
32 {"Bedroom": 1, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1570168, 42.3472801]}, "properties":
33 {"Bedroom": 2, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1034359, 42.3884241]}, "properties":
34 {"Bedroom": 2, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.037249, 42.3297091]}, "properties":
35 {"Bedroom": 2, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1157314, 42.2897066]}, "properties":
36 {"Bedroom": 3, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1319121, 42.3558468]}, "properties":
37 {"Bedroom": 3, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1495, 42.2883]}, "properties":
38 {"Bedroom": 3, "Price": 900}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.1714369, 42.4212024]}, "properties":
39 {"Bedroom": 0, "Price": 925}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.042408, 42.388433]}, "properties":
40 {"Bedroom": 0, "Price": 925}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0424, 42.3884]}, "properties":
41 {"Bedroom": 0, "Price": 925}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0931052, 42.3487795]}, "properties":
42 {"Bedroom": 0, "Price": 925}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.099896, 42.367876]}, "properties":
43 {"Bedroom": 1, "Price": 925}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.0931052, 42.3487795]}, "properties":
44 {"Bedroom": 1, "Price": 925}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.143445, 42.4062898]}, "properties":
45 {"Bedroom": 0, "Price": 950}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-71.069324, 42.3484359]}, "properties":
46 {"Bedroom": 0, "Price": 950}, {"type": "Feature", "geometry": {"type": "Point", "coordinates": [-70.9911973, 42.4186969]}, "properties":
```

# 3 Kinds of Dataset Sources

- (3) Manually Collected Data

In order to get first-hand detailed rental information for training models, we collected data of different areas in Boston from <http://www.apartments.com> and <http://www.zillow.com>.

Parameters: Zip-code, number of bedroom, number of bathroom, building type(house/apartment/condo/townhome), 24-emergency/maintenance(yes/no), GYM, AC, free-water, free-electricity, free-heat, dishwasher, microwave, washer-dryer, public-transport, price

free  
was  
nea

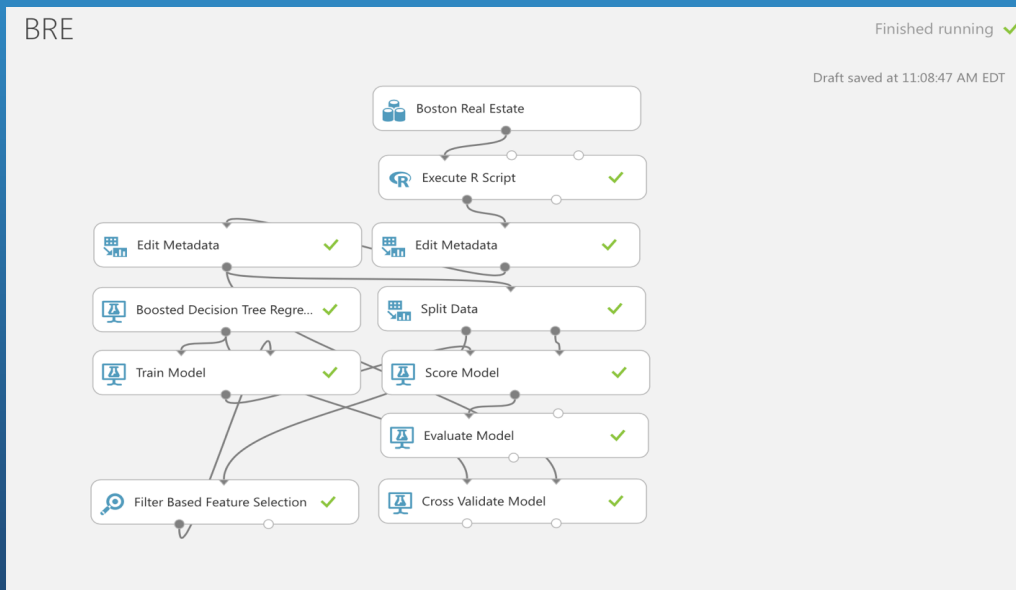
- Data

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
zipcode	bedroom	bathroom	type	parking	24h-emerg	GYM	AC	free-water	free-electric	free-heat	dishwasher	microwave	washer-dryer	public-transp	price
2108	1	1	2	1	1	1	1	1	0	1	1	1	1	1	2200
2108	2	1	4	0	0	0	0	1	0	1	1	1	2	1	1400
2108	2	2	2	2	1	1	1	1	0	1	1	1	1	2	5500
2108	2	2.5	2	2	1	1	1	1	0	1	1	1	1	2	7000
2108	2	2	2	2	0	0	1	1	0	1	1	1	2	1	1030
2108	2	2	2	2	1	0	1	1	0	1	1	1	1	1	6600
2108	1	1	3	2	0	0	0	1	0	1	0	0	2	1	1450
2108	1	1	2	0	0	0	0	1	0	1	1	0	1	1	3050
2108	1	1	2	0	0	0	0	1	0	1	1	0	1	1	3100
2108	1	1.5	2	1	1	1	1	1	0	1	1	1	1	1	4000
2108	2	2.5	2	1	1	1	1	1	0	1	1	1	1	1	6800
2108	1	1	2	1	1	1	1	1	0	1	1	1	2	1	2200
2108	2	1	4	0	0	0	0	1	0	1	0	1	1	1	1400
2108	2	3	3	2	0	0	0	1	0	1	1	1	0	1	6200
2108	1	1	2	2	0	0	0	1	0	1	1	0	2	1	2600
2108	2	1	3	0	0	0	0	1	0	1	1	1	2	1	2900
2108	1	1.5	3	1	1	1	1	1	1	1	1	1	1	1	4500
2108	1	1	1	0	0	0	1	1	0	1	1	0	2	1	2900
2109	1	1	2	1	1	1	1	1	0	1	1	1	1	1	3025
2109	2	2	2	1	1	1	1	1	0	1	1	1	1	1	4420
2109	1	1	2	1	1	1	1	1	0	1	1	1	1	1	3800
2109	2	1	2	1	1	1	1	1	0	1	1	1	1	1	4000

# 3 Tools We Used:

- (1) Azure Studio

We compared different regression models and deployed 2 web service. One is for quick prediction in our web application, the other is for advanced prediction.



# 3 Tools We Used:

- Azure Studio

Price ~ zip + bedroom + bathroom + type					
	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error	Relative Squared Error	Coefficient of determination( $R^2$ )
Boosted Decision Tree Regression	575	888	Price ~ zip + bedroom + bathroom + type + 24-emergency/maintenance + GYM + AC + free-water + free-electricity + free-heat + dishwasher + microwave + washer-dryer + public-transportation		
Bayesian Linear Regression	616	899			
Decision Forest Regression	619	967	Boosted Decision Tree Regression	556	841
Linear Regression	613	891	Bayesian Linear Regression	591	866
Neural Network Regression	1777	2061	Decision Forest Regression	602	918
Poisson Regression	769	1080	Linear Regression	498	878
Fast Forest Quantile Regression	Average Quantile Loss: 318		Neural Network Regression	1770	2054
			Poisson Regression	657	960
			Fast Forest Quantile Regression	Average Quantile Loss: 286	

s deployed 2  
n our web

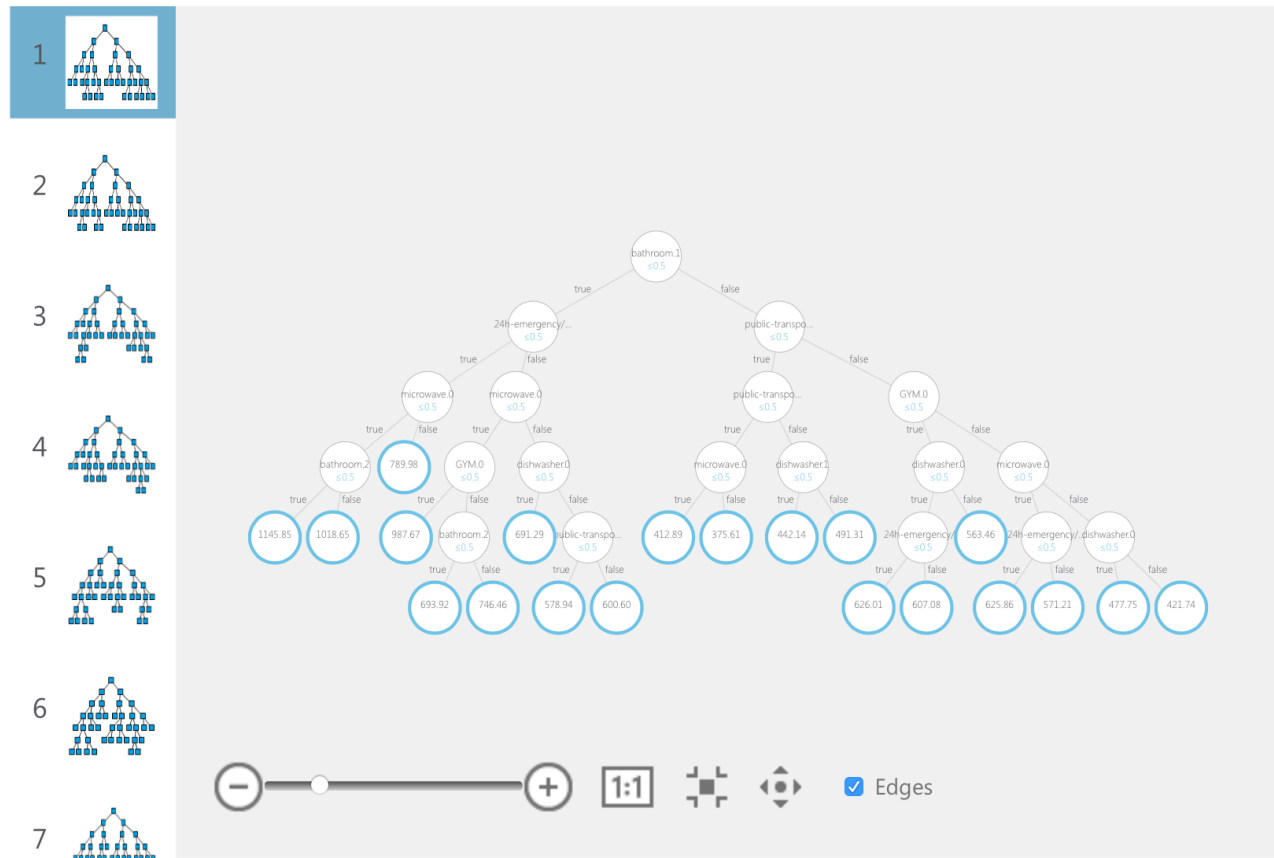


# 3 Tools We Used:

- Azure Machine Learning - Features > Train Model > Trained model

Price ~ trees constructed  
100

Boosted  
Decision  
Tree  
Regression  
Bayesian  
Linear  
Regression  
Decision  
Forest  
Regression  
Linear  
Regression  
Neural Network  
Regression  
Poisson  
Regression  
Fast Forest  
Quantile  
Regression



nce +

cient of  
mination(

Regression

Fast  
Forest  
Quantile

Average  
Quantile  
Loss: 286

# 3 Tools We Used:

- (2) Azure Portal

We stored the geojson files we mentioned before using

The screenshot displays the Microsoft Azure Portal interface for a storage account named 'finalprojbostonrent'. The interface is divided into several sections:

- Monitoring:** Located on the left, it shows a 'Total requests' chart and a table of metrics. The table has four columns: BLOB (4), TABLE (938), QUEUE (4), and FILE (2.46K). Below this, there are three more charts: 'Total egress', 'Average E2E latency', and 'Success'.
- Navigation:** A central sidebar lists various Azure services: Resource groups, All resources, Recent, App Services, Virtual machines (classic), Virtual machines, SQL databases, Cloud services (classic), Security Center, and Subscriptions. A 'More Services' link is at the bottom.
- Overview:** The main section on the right, titled 'finalprojbostonrent Storage account - General'. It includes a search bar and a list of links: Overview (selected), Activity log, Access control (IAM), Tags, and Diagnose and solve problems. Below this is a 'SETTINGS' section with links to Access keys, Configuration, Custom domain, Encryption, and Shared access signature.
- Essentials:** A section on the right showing account details: Resource group (finalprojresourcegroup), Status (Primary: Available, Secondary: Available), Location (Central US, East US 2), Subscription name (免费试用), and Subscription ID (70a7d39e-34b9-47b5-a59b-2fa34ccdc8f). It also shows Performance (Standard) and Replication (Read-access geo-redundant storage (RA-G...)).
- Essentials Tiles:** A section below Essentials with four tiles: Blobs (highlighted with a blue box), Files, Tables, and Queues. Each tile has an 'Add tiles' button.
- Monitoring Tiles:** A section at the bottom with a 'Total requests' tile and an 'Add tiles' button.

# 3 Tools We Used:

- (2) Azure Portal

Microsoft Azure finalprojbostonrent > File service > asd > Upload files

File service (finalprojbostonrent)

Settings File share Refresh

Essentials

Search file shares by prefix

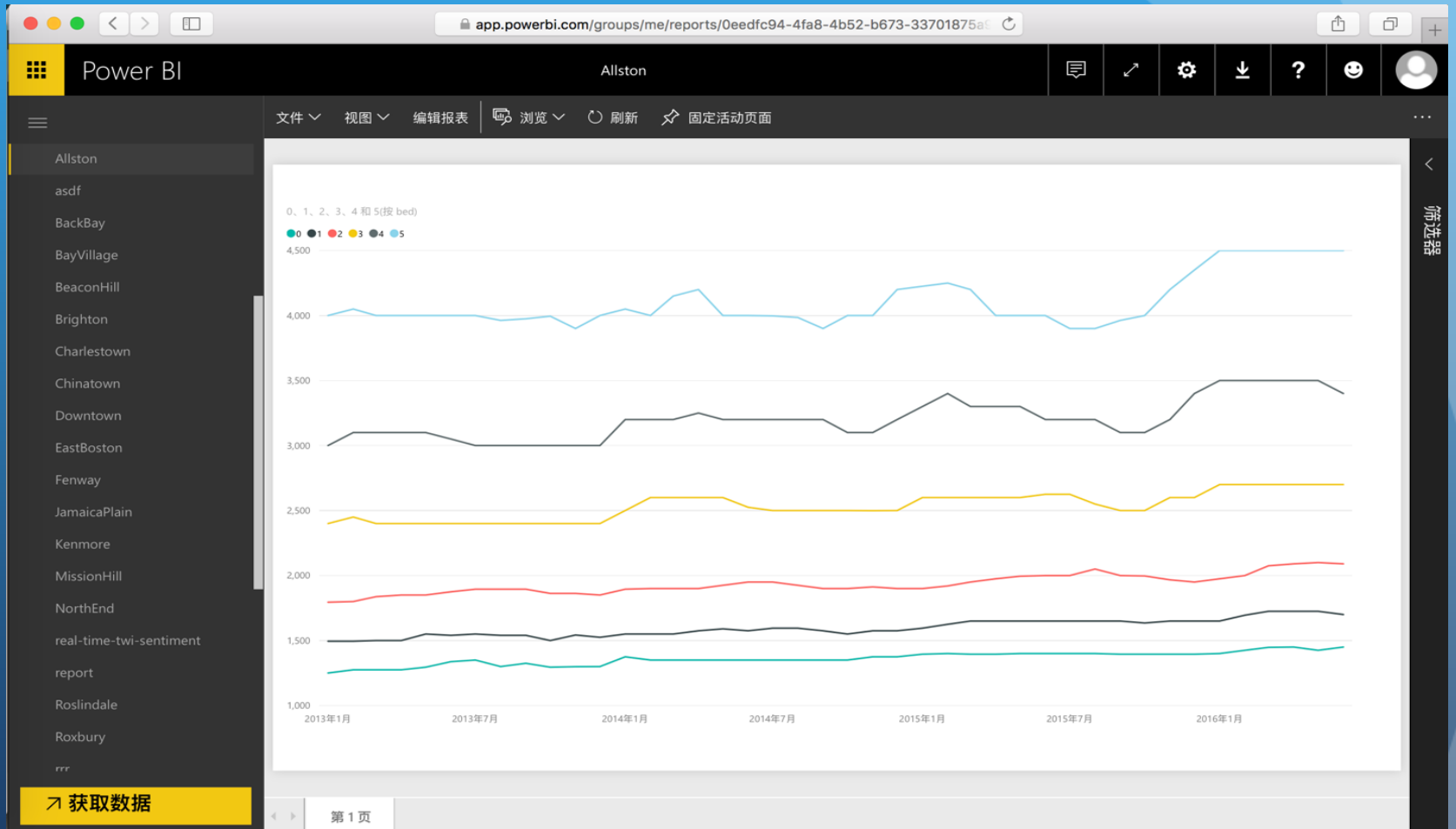
NAME	MODIFIED	QUOTA
asd	8/18/2016, 8:06:30 PM	5 TB

Connect Upload Directory Refresh Delete

NAME	TYPE
2013-02-18.geojson	File
2013-03-18.geojson	File
2013-04-18.geojson	File
2013-05-18.geojson	File
2013-06-18.geojson	File
2013-07-18.geojson	File
2013-08-18.geojson	File
2013-09-18.geojson	File
2013-10-18.geojson	File
2013-11-18.geojson	File
2013-11-21.geojson	File
2013-12-18.geojson	File
20130.geojson	File
20131.geojson	File
20132.geojson	File

# 3 Tools We Used:

- (3) Power BI



# Web application

<http://adsfinalproject.xigtazbqwr.us-east-1.elasticbeanstalk.com/>

