In [1]:

```
import pandas as pd
import numpy as np
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns
```

In [39]:

```
pizza_data = pd.read_csv("C:/Users/My PC/Downloads/Python/pizza/8358_1.csv")
```

In [40]:

```
pizza_data.head()
```

Out[40]:

	id	address	categories	city	country	
0	AVwc_6KEIN2L1WUfrKAH	Cascade Village Mall Across From Target	Pizza Place	Bend	US	us/or/bend/cas
1	AVwc_6KEIN2L1WUfrKAH	Cascade Village Mall Across From Target	Pizza Place	Bend	US	us/or/bend/cas
2	AVwc_6qRByjofQCxkcxw	148 S Barrington Ave	American Restaurant,Bar,Bakery	Los Angeles	US	us/brentwood/los
3	AVwc_6qRByjofQCxkcxw	148 S Barrington Ave	American Restaurant,Bar,Bakery	Los Angeles	US	us/brentwood/los
4	AVwc_6qRByjofQCxkcxw	148 S Barrington Ave	American Restaurant,Bar,Bakery	Los Angeles	US	us/brentwood/los

5 rows × 21 columns

In [41]:

```
pizza_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3510 entries, 0 to 3509
Data columns (total 21 columns):
id                    3510 non-null object
address               3510 non-null object
categories            3510 non-null object
city                  3510 non-null object
country               3510 non-null object
keys                  3510 non-null object
latitude              3510 non-null float64
longitude             3510 non-null float64
menuPageURL           610 non-null object
menus.amountMax       2948 non-null float64
menus.amountMin       2948 non-null float64
menus.currency        2951 non-null object
menus.dateSeen        3510 non-null object
menus.description     1526 non-null object
menus.name            3510 non-null object
name                  3510 non-null object
postalCode            3484 non-null object
priceRangeCurrency    1557 non-null object
priceRangeMin         1557 non-null float64
priceRangeMax         1557 non-null float64
province              3510 non-null object
dtypes: float64(6), object(15)
memory usage: 576.0+ KB
```

In [42]:

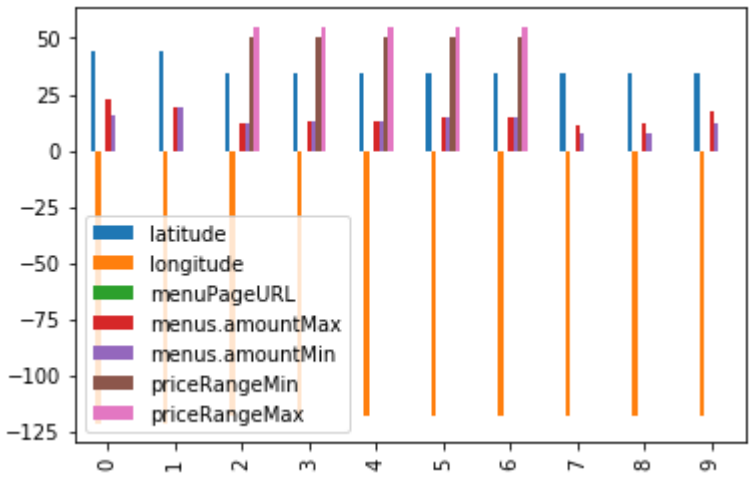
```
pizzas = pizza_data['menus.name'].value_counts()
pizzas
```

Out[42]:

```
Cheese Pizza                          136
White Pizza                            75
Margherita Pizza                       63
Pizza                                  62
Hawaiian Pizza                         54
                                     ...
Bolognia Pizza                          1
Mushroom Pizza Steak                    1
The Tommy Gun Breakfast Pizza           1
Any Six ( 6 ) 10 Signature Pizzas       1
Pizza Parmigiana                        1
Name: menus.name, Length: 1598, dtype: int64
```

In [43]:

```
pizzas = pizza_data[:10].plot.bar()
```



In [44]:

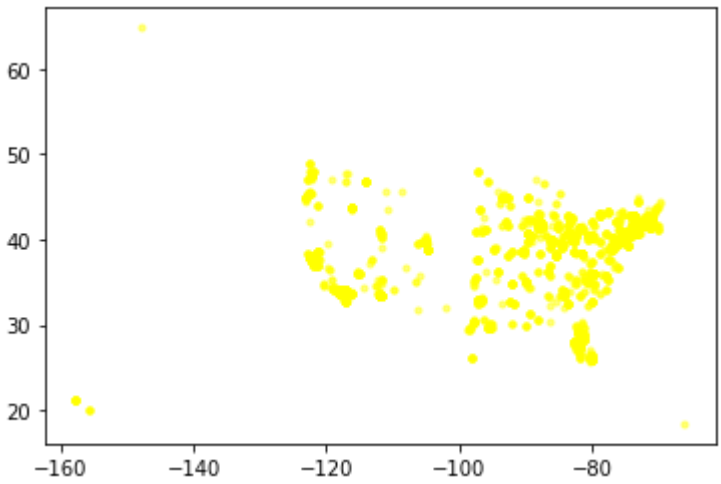
```
longitude = (min(pizza_data['longitude']),max(pizza_data['longitude']))
```

In [45]:

```
latitude = (min(pizza_data['latitude']),max(pizza_data['latitude']))
```

In [46]:

```
xy = plt.scatter(pizza_data['longitude'].values,pizza_data['latitude'].values,color='ye
llow',s=10,alpha=0.5)
plt.show()
```



In [47]:

```
pizzas_city = pizza_data['city'].value_counts()
pizzas_city
```

Out[47]:

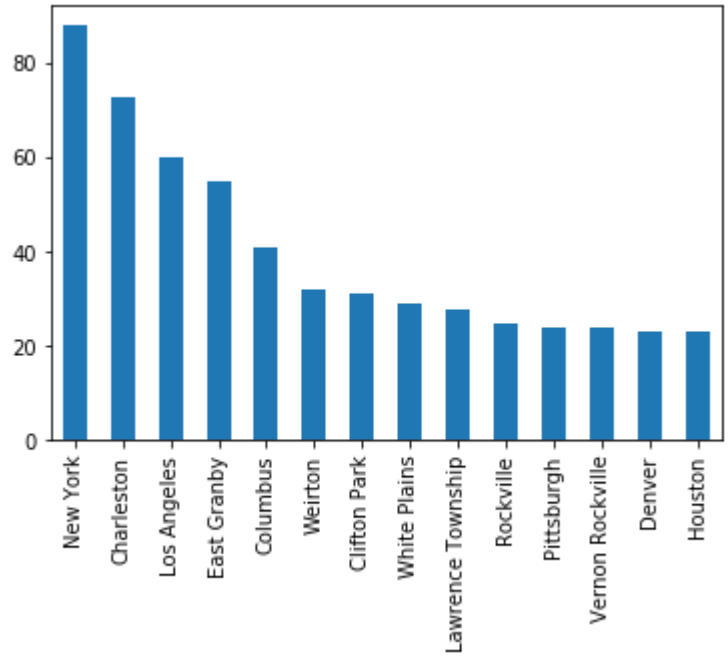
```
Philadelphia     91
New York         88
Charleston       73
Los Angeles      60
East Granby      55
                 ..
Monroeville       1
Waterville        1
New Brunswick     1
Sartell           1
Santa Fe          1
Name: city, Length: 673, dtype: int64
```

In [48]:

```
pizzas_city[1:15].plot.bar()
```

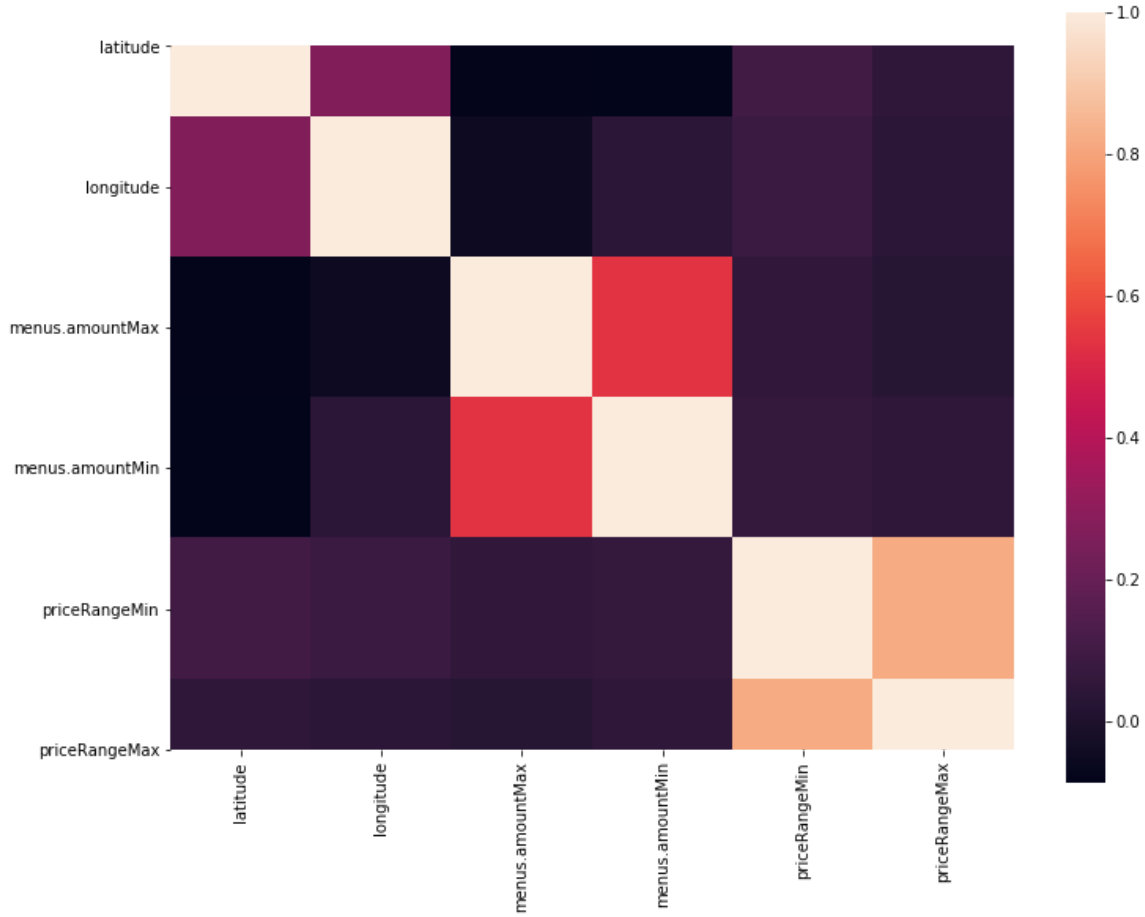
Out[48]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b3c3cdf448>
```



In [49]:

```
f, ax = plt.subplots(figsize=(12, 9))
sns.heatmap(pizza_data.corr(), square=True);
```



In [50]:

```
menupizza = pizza_data[['menus.amountMax', 'menus.amountMin']]
menupizza.describe()
```

Out[50]:

	menus.amountMax	menus.amountMin
count	2948.000000	2948.000000
mean	12.479186	11.427849
std	13.398923	7.262919
min	0.000000	0.000000
25%	7.500000	7.000000
50%	11.870000	10.990000
75%	15.950000	14.950000
max	312.950000	118.990000

In [51]:

```
pricepizza = pizza_data[['priceRangeMax', 'priceRangeMin']]
pricepizza.describe()
```

Out[51]:

	priceRangeMax	priceRangeMin
count	1557.000000	1557.000000
mean	36.566474	15.597945
std	21.737839	18.495854
min	12.000000	0.000000
25%	30.000000	0.000000
50%	40.000000	25.000000
75%	40.000000	25.000000
max	666.000000	347.000000

In [52]:

```
citydata = pizza_data['city'].value_counts().reset_index()
citydata.columns = ['city', 'values']
citydata
```

Out[52]:

	city	values
0	Philadelphia	91
1	New York	88
2	Charleston	73
3	Los Angeles	60
4	East Granby	55
...	...	...
668	Monroeville	1
669	Waterville	1
670	New Brunswick	1
671	Sartell	1
672	Santa Fe	1

673 rows × 2 columns

In [53]:

```
pizza = pd.read_csv("C:/Users/My PC/Downloads/Python/pizza/Datafiniti_Pizza_Restaurants_and_the_Pizza_They_Sell_May19.csv")
```

In [54]:

```
pizza.head()
```

Out[54]:

	id	dateAdded	dateUpdated	address	categories
0	AVz3Y-7h3D1zeR_xDAqm	2017-06-30T05:05:40Z	2019-05-01T15:43:09Z	4203 E Kiehl Ave	Pizza,Restaurant,American restaurants,Pizza Pl...
1	AVweGPFF_7pvs4fzAAzQ	2016-04-02T04:02:49Z	2019-05-01T15:27:50Z	25 E Camelback Rd	Pizza,Pizza Place,Restaurants
2	AVwdRGa9_7pvs4fz4E3K	2016-03-03T18:39:49Z	2019-05-01T12:52:25Z	3703 Paxton Ave	Restaurant,Pizza Place,Restaurants
3	AVwdX4psIN2L1WUfvJB1	2016-03-29T05:08:59Z	2019-05-01T12:52:20Z	30495 John R Rd	Pizza,Carry-out food,Pizza Place,Restaurants
4	AVwdaeTtkufWRAb55pSH	2016-03-31T02:34:04Z	2019-05-01T12:50:45Z	3600 Eastern Ave	Pizza,American restaurants,Pizza Place,Pizza e...

5 rows × 24 columns

In [65]:

```
pzs = pizza['menus.name'].value_counts()
pzs
```

Out[65]:

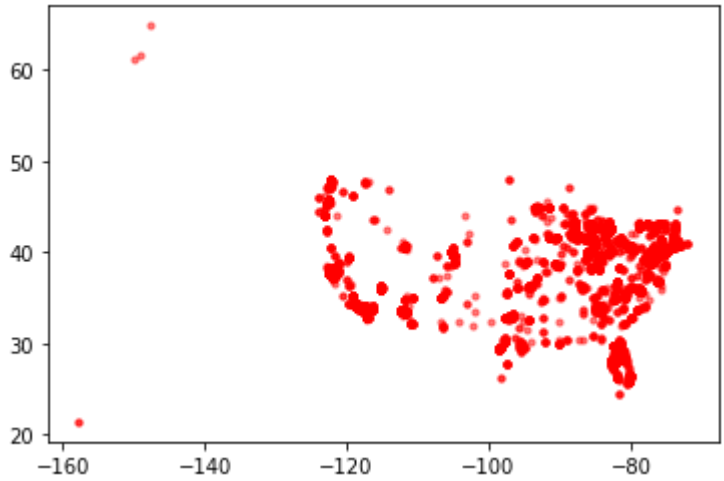
```
Cheese Pizza                    299
White Pizza                     225
Hawaiian Pizza                  149
Buffalo Chicken Pizza           112
Pizza                           111
                               ...
18ampquot Pizza Grinder           1
Ranchero Pizza With Chicken       1
Veggie Stuffed Pizza              1
Spaghetti Pizza                   1
White Pizza (3 Cheeses)           1
Name: menus.name, Length: 4749, dtype: int64
```

In [71]:

```
lat = (min(pizza['latitude']),max(pizza['latitude']))
long = (min(pizza['longitude']),max(pizza['longitude']))
```

In [72]:

```
yx = plt.scatter(pizza['longitude'].values,pizza['latitude'].values,color='red',s=10,al
pha=0.5)
plt.show()
```



In [76]:

```
menupizza = pizza[['menus.amountMax', 'menus.amountMin']]
menupizza.describe()
```

Out[76]:

	menus.amountMax	menus.amountMin
count	10000.000000	10000.000000
mean	14.032670	13.416902
std	17.607233	10.742678
min	0.000000	0.000000
25%	8.490000	8.000000
50%	13.390000	12.990000
75%	17.950000	16.990000
max	1395.000000	243.000000

In [75]:

```
pricepizza = pizza[['priceRangeMax', 'priceRangeMin']]
pricepizza.describe()
```

Out[75]:

	priceRangeMax	priceRangeMin
count	10000.000000	10000.000000
mean	27.764200	4.655500
std	6.017526	9.828412
min	7.000000	0.000000
25%	25.000000	0.000000
50%	25.000000	0.000000
75%	25.000000	0.000000
max	55.000000	50.000000

In [79]:

```
citydata = pizza['city'].value_counts().reset_index()
citydata.columns = ['city', 'values']
citydata
```

Out[79]:

	city	values
0	New York	655
1	Brooklyn	460
2	Los Angeles	193
3	Buffalo	178
4	Philadelphia	140
...	...	...
1023	Alma	1
1024	Urbana	1
1025	Glen Mills	1
1026	Kittanning	1
1027	Odessa	1

1028 rows × 2 columns

In [78]:

```
pizzas_city = pizza['city'].value_counts()
pizzas_city
```

Out[78]:

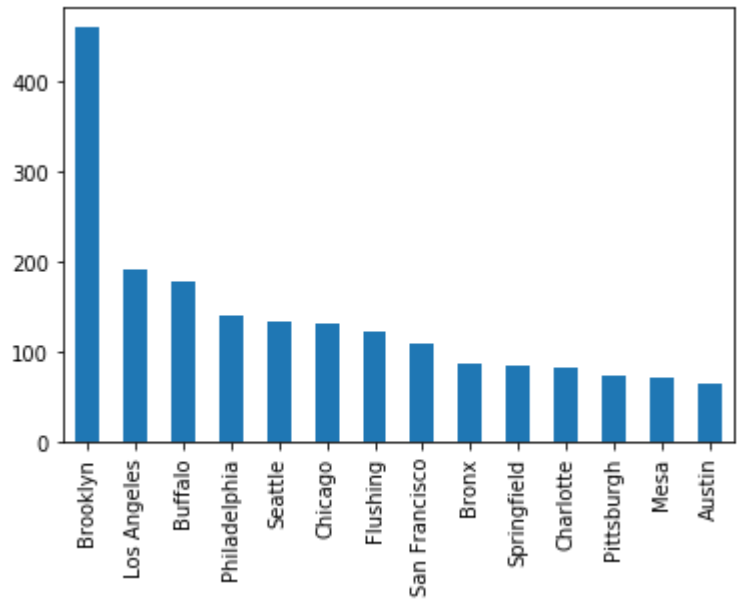
```
New York        655
Brooklyn        460
Los Angeles     193
Buffalo         178
Philadelphia    140
               ...
Alma              1
Urbana            1
Glen Mills        1
Kittanning        1
Odessa            1
Name: city, Length: 1028, dtype: int64
```

In [80]:

```
pizzas_city[1:15].plot.bar()
```

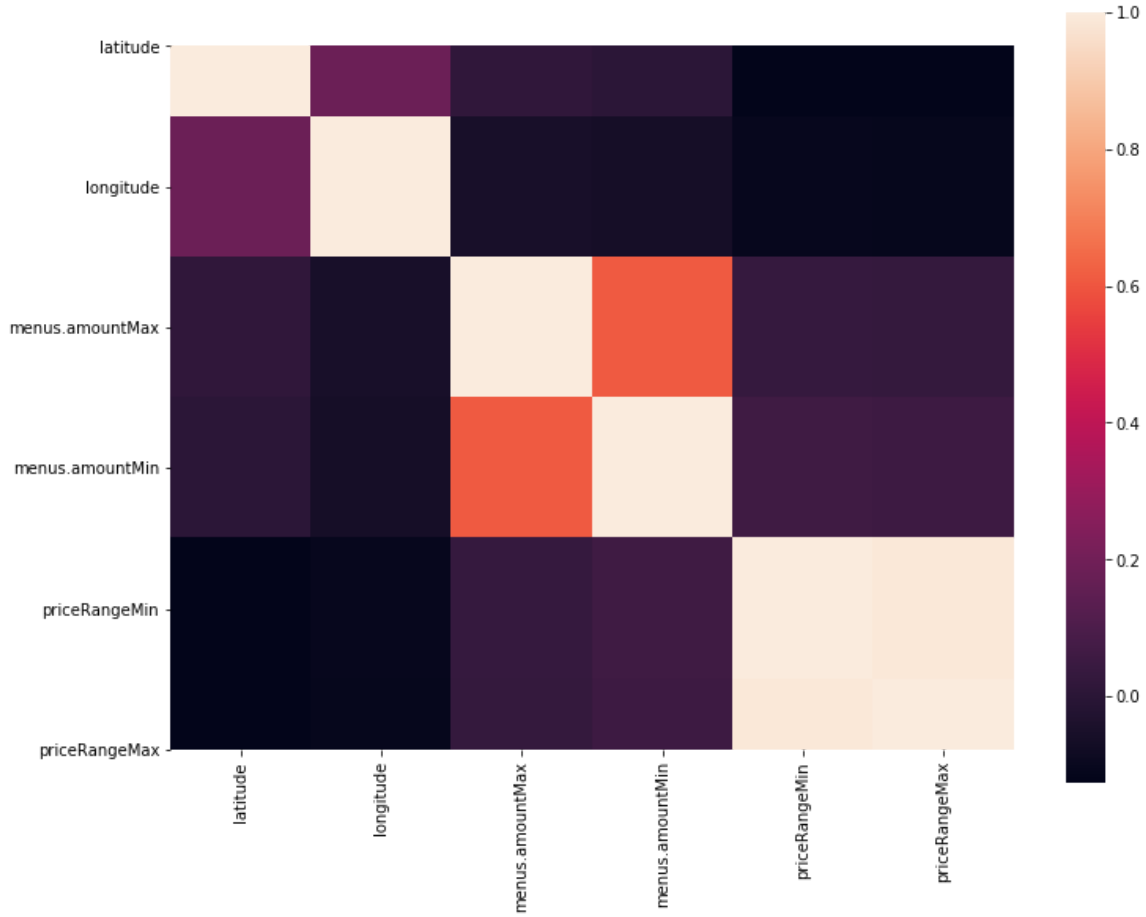
Out[80]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1b3c1f33208>
```



In [82]:

```
f, ax = plt.subplots(figsize=(12, 9))
sns.heatmap(pizza.corr(), square=True);
```



In [ ]:

```

```