chipotle 데이터

탐색

In [1]:
```
import pandas as pd
```

In [2]:
```
# data 불러오기
# 이때 데이터를 Dataframe 형태로 불러옴.
chipo = pd.read_csv("C:/Users/sungwonKim/Desktop/SAM/이것이 데이터분석이다 with 파이썬
```

In [3]:
```
chipo.head()
```

Out[3]:

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Chips and Fresh Tomato Salsa	NaN	$2.39
1	1	1	Izze	[Clementine]	$3.39
2	1	1	Nantucket Nectar	[Apple]	$3.39
3	1	1	Chips and Tomatillo-Green Chili Salsa	NaN	$2.39
4	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	$16.98

In [4]:
```
# id : 송장 번호 개념 / quantity : 개수 / item_name : 뭘 시켰는지?
# choice_description : 토핑 / item_price : 총 금액
```

In [5]:
```
print(chipo.shape)
print('------------------------')
print(chipo.info())
```

```
(4622, 5)
------------------------
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype 
---  ------              --------------  ----- 
 0   order_id            4622 non-null   int64 
 1   quantity            4622 non-null   int64 
 2   item_name           4622 non-null   object
 3   choice_description  3376 non-null   object
 4   item_price          4622 non-null   object
dtypes: int64(2), object(3)
memory usage: 180.7+ KB
None
```

In [6]:
```
# order_id , quantity : int
# item_name, choice_description, item_price : str 형태
```

In [ ]:
```

```

In [7]:
```
# chipo 라는 Dataframe에서 순서대로 10개의 row 데이터를 show
chipo.head(10)
```

Out[7]:

	order_id	quantity	item_name	choice_description	item_price
0	1	1	Chips and Fresh Tomato Salsa	NaN	$2.39
1	1	1	Izze	[Clementine]	$3.39
2	1	1	Nantucket Nectar	[Apple]	$3.39
3	1	1	Chips and Tomatillo-Green Chili Salsa	NaN	$2.39
4	2	2	Chicken Bowl	[Tomatillo-Red Chili Salsa (Hot), [Black Beans...	$16.98
5	3	1	Chicken Bowl	[Fresh Tomato Salsa (Mild), [Rice, Cheese, Sou...	$10.98
6	3	1	Side of Chips	NaN	$1.69
7	4	1	Steak Burrito	[Tomatillo Red Chili Salsa, [Fajita Vegetables...	$11.75
8	4	1	Steak Soft Tacos	[Tomatillo Green Chili Salsa, [Pinto Beans, Ch...	$9.25
9	5	1	Steak Burrito	[Fresh Tomato Salsa, [Rice, Black Beans, Pinto...	$9.25

In [8]:
```
print(chipo.columns)
print('---------------------------')
print(chipo.index)
```

```
Index(['order_id', 'quantity', 'item_name', 'choice_description',
       'item_price'],
      dtype='object')
---------------------------
RangeIndex(start=0, stop=4622, step=1)
```

In [9]:
```
chipo['order_id'] = chipo['order_id'].astype(str)
#order_id는 숫자의 의미가 없다. 그래서 str로 변환.
```

In [10]:
```
print(chipo.describe())
# chipo dataframe에서 수치형 피처들의 요약 통계량을 확인
```

```
          quantity
count  4622.000000
mean      1.075725
std       0.410186
min       1.000000
25%       1.000000
50%       1.000000
75%       1.000000
max      15.000000
```

In [11]:
```
# order_id의 개수 출력
print(len(chipo['order_id'].unique()))
```

```
1834
```

In [12]:
```
# item_name(메뉴)의 개수 출력
```

```
print(len(chipo['item_name'].unique()))
```

```
50
```

In [13]:
```
# 수치형 : describe
# 범주형 : unique (item name 이 총 50개 있음을 보여줬음.)
```

In [14]:
```
chipo['item_name'].value_counts()
```

Out[14]:
```
Chicken Bowl                            726
Chicken Burrito                         553
Chips and Guacamole                     479
Steak Burrito                           368
Canned Soft Drink                       301
Chips                                   211
Steak Bowl                              211
Bottled Water                           162
Chicken Soft Tacos                      115
Chicken Salad Bowl                      110
Chips and Fresh Tomato Salsa            110
Canned Soda                             104
Side of Chips                           101
Veggie Burrito                           95
Barbacoa Burrito                         91
Veggie Bowl                              85
Carnitas Bowl                            68
Barbacoa Bowl                            66
Carnitas Burrito                         59
Steak Soft Tacos                         55
6 Pack Soft Drink                        54
Chips and Tomatillo Red Chili Salsa      48
Chicken Crispy Tacos                     47
Chips and Tomatillo Green Chili Salsa    43
Carnitas Soft Tacos                      40
Steak Crispy Tacos                       35
Chips and Tomatillo-Green Chili Salsa    31
Steak Salad Bowl                         29
Nantucket Nectar                         27
Barbacoa Soft Tacos                      25
Chips and Roasted Chili Corn Salsa       22
Izze                                     20
Chips and Tomatillo-Red Chili Salsa      20
Veggie Salad Bowl                        18
Chips and Roasted Chili-Corn Salsa       18
Barbacoa Crispy Tacos                    11
Barbacoa Salad Bowl                      10
Chicken Salad                             9
Veggie Soft Tacos                         7
Carnitas Crispy Tacos                     7
Veggie Salad                              6
Burrito                                   6
Carnitas Salad Bowl                       6
Steak Salad                               4
Bowl                                      2
Crispy Tacos                              2
Salad                                     2
Veggie Crispy Tacos                       1
Chips and Mild Fresh Tomato Salsa         1
Carnitas Salad                            1
Name: item_name, dtype: int64
```

In [15]:
```
chipo['item_name'].value_counts().sum()
```

Out[15]:
```
4622
```

In [16]:
```
# value counts 는 각 item_name 별로 몇 개가 있는지 보여줌.
```

In [17]:
```
# 길이를 뜻하는 len을 함수를 이용하면 unique와 같은 결과 나타냄.
```

In [18]:
```
len(chipo['item_name'].value_counts())
```

Out[18]: 50

탐색 및 시각화

탐색

In [19]:
```
# 가장 많이 주문한 item : top 10 출력
item_count = chipo['item_name'].value_counts()[:10]
for idx, (val, cnt) in enumerate(item_count.iteritems(), 1) :
    print("Top", idx, ":", val, cnt)
```

```
Top 1 : Chicken Bowl 726
Top 2 : Chicken Burrito 553
Top 3 : Chips and Guacamole 479
Top 4 : Steak Burrito 368
Top 5 : Canned Soft Drink 301
Top 6 : Chips 211
Top 7 : Steak Bowl 211
Top 8 : Bottled Water 162
Top 9 : Chicken Soft Tacos 115
Top 10 : Chicken Salad Bowl 110
```

In [20]:
```
chipo['item_name'].value_counts()[0:10]
### Question : [0:10] 에서 10이 의미하는 게 무엇인가요? ###
```

Out[20]:
```
Chicken Bowl            726
Chicken Burrito         553
Chips and Guacamole     479
Steak Burrito           368
Canned Soft Drink       301
Chips                   211
Steak Bowl              211
Bottled Water           162
Chicken Soft Tacos      115
Chicken Salad Bowl      110
Name: item_name, dtype: int64
```

In [21]:
```
chipo['item_name']. value_counts()
```

Out[21]:
```
Chicken Bowl                          726
Chicken Burrito                       553
Chips and Guacamole                   479
Steak Burrito                         368
Canned Soft Drink                     301
Chips                                 211
Steak Bowl                            211
Bottled Water                         162
Chicken Soft Tacos                    115
Chicken Salad Bowl                    110
Chips and Fresh Tomato Salsa          110
Canned Soda                           104
Side of Chips                         101
Veggie Burrito                         95
```

```
Barbacoa Burrito                        91
Veggie Bowl                             85
Carnitas Bowl                           68
Barbacoa Bowl                           66
Carnitas Burrito                        59
Steak Soft Tacos                        55
6 Pack Soft Drink                       54
Chips and Tomatillo Red Chili Salsa     48
Chicken Crispy Tacos                    47
Chips and Tomatillo Green Chili Salsa   43
Carnitas Soft Tacos                     40
Steak Crispy Tacos                      35
Chips and Tomatillo-Green Chili Salsa   31
Steak Salad Bowl                        29
Nantucket Nectar                        27
Barbacoa Soft Tacos                     25
Chips and Roasted Chili Corn Salsa      22
Izze                                    20
Chips and Tomatillo-Red Chili Salsa     20
Veggie Salad Bowl                       18
Chips and Roasted Chili-Corn Salsa      18
Barbacoa Crispy Tacos                   11
Barbacoa Salad Bowl                     10
Chicken Salad                            9
Veggie Soft Tacos                        7
Carnitas Crispy Tacos                    7
Veggie Salad                             6
Burrito                                  6
Carnitas Salad Bowl                      6
Steak Salad                              4
Bowl                                     2
Crispy Tacos                             2
Salad                                    2
Veggie Crispy Tacos                      1
Chips and Mild Fresh Tomato Salsa        1
Carnitas Salad                           1
Name: item_name, dtype: int64
```

In [22]:
```
chipo['item_name'].value_counts().index[0]
```

Out[22]:
```
'Chicken Bowl'
```

In [23]:
```
chipo['item_name'].value_counts().tolist()[0]
```

Out[23]:
```
726
```

In [24]:
```
chipo['item_name'].value_counts().index.tolist()[0]
```

Out[24]:
```
'Chicken Bowl'
```

In [25]:
```
# item당 주문 개수 (송장 번호 개수의 의미)
order_count = chipo.groupby('item_name')['order_id'].count()
# item_name 을 토대로 group by / 지켜보는 걸 order_id
order_count[:10]
```

Out[25]:
```
item_name
6 Pack Soft Drink         54
Barbacoa Bowl             66
Barbacoa Burrito          91
Barbacoa Crispy Tacos     11
Barbacoa Salad Bowl       10
Barbacoa Soft Tacos       25
```

```
Bottled Water            162
Bowl                       2
Burrito                    6
Canned Soda              104
Name: order_id, dtype: int64
```

In [26]:
```
item_quantity = chipo.groupby('item_name')['quantity'].sum()
item_quantity
```

Out[26]:
```
item_name
6 Pack Soft Drink                         55
Barbacoa Bowl                             66
Barbacoa Burrito                          91
Barbacoa Crispy Tacos                     12
Barbacoa Salad Bowl                       10
Barbacoa Soft Tacos                       25
Bottled Water                            211
Bowl                                       4
Burrito                                    6
Canned Soda                              126
Canned Soft Drink                        351
Carnitas Bowl                             71
Carnitas Burrito                          60
Carnitas Crispy Tacos                      8
Carnitas Salad                             1
Carnitas Salad Bowl                        6
Carnitas Soft Tacos                       40
Chicken Bowl                             761
Chicken Burrito                          591
Chicken Crispy Tacos                      50
Chicken Salad                              9
Chicken Salad Bowl                       123
Chicken Soft Tacos                       120
Chips                                    230
Chips and Fresh Tomato Salsa             130
Chips and Guacamole                      506
Chips and Mild Fresh Tomato Salsa          1
Chips and Roasted Chili Corn Salsa        23
Chips and Roasted Chili-Corn Salsa        18
Chips and Tomatillo Green Chili Salsa     45
Chips and Tomatillo Red Chili Salsa       50
Chips and Tomatillo-Green Chili Salsa     33
Chips and Tomatillo-Red Chili Salsa       25
Crispy Tacos                               2
Izze                                      20
Nantucket Nectar                          29
Salad                                      2
Side of Chips                            110
Steak Bowl                               221
Steak Burrito                            386
Steak Crispy Tacos                        36
Steak Salad                                4
Steak Salad Bowl                          31
Steak Soft Tacos                          56
Veggie Bowl                               87
Veggie Burrito                            97
Veggie Crispy Tacos                        1
Veggie Salad                               6
Veggie Salad Bowl                         18
Veggie Soft Tacos                          8
Name: quantity, dtype: int64
```

In [27]:
```
# item당 주문 총량 (한 사람이 2팩이상 사는 경우가 있기에 다르다.)
item_quantity = chipo.groupby('item_name')['quantity'].sum()
item_quantity[:10]
```

```
item_name
```

Out[27]:
```
6 Pack Soft Drink         55
Barbacoa Bowl             66
Barbacoa Burrito          91
Barbacoa Crispy Tacos     12
Barbacoa Salad Bowl       10
Barbacoa Soft Tacos       25
Bottled Water            211
Bowl                       4
Burrito                    6
Canned Soda              126
Name: quantity, dtype: int64
```

시각화

In [28]:
```
item_quantity.index.tolist()
```

Out[28]:
```
['6 Pack Soft Drink',
 'Barbacoa Bowl',
 'Barbacoa Burrito',
 'Barbacoa Crispy Tacos',
 'Barbacoa Salad Bowl',
 'Barbacoa Soft Tacos',
 'Bottled Water',
 'Bowl',
 'Burrito',
 'Canned Soda',
 'Canned Soft Drink',
 'Carnitas Bowl',
 'Carnitas Burrito',
 'Carnitas Crispy Tacos',
 'Carnitas Salad',
 'Carnitas Salad Bowl',
 'Carnitas Soft Tacos',
 'Chicken Bowl',
 'Chicken Burrito',
 'Chicken Crispy Tacos',
 'Chicken Salad',
 'Chicken Salad Bowl',
 'Chicken Soft Tacos',
 'Chips',
 'Chips and Fresh Tomato Salsa',
 'Chips and Guacamole',
 'Chips and Mild Fresh Tomato Salsa',
 'Chips and Roasted Chili Corn Salsa',
 'Chips and Roasted Chili-Corn Salsa',
 'Chips and Tomatillo Green Chili Salsa',
 'Chips and Tomatillo Red Chili Salsa',
 'Chips and Tomatillo-Green Chili Salsa',
 'Chips and Tomatillo-Red Chili Salsa',
 'Crispy Tacos',
 'Izze',
 'Nantucket Nectar',
 'Salad',
 'Side of Chips',
 'Steak Bowl',
 'Steak Burrito',
 'Steak Crispy Tacos',
 'Steak Salad',
 'Steak Salad Bowl',
 'Steak Soft Tacos',
 'Veggie Bowl',
 'Veggie Burrito',
 'Veggie Crispy Tacos',
 'Veggie Salad',
 'Veggie Salad Bowl',
 'Veggie Soft Tacos']
```

In [29]:

```
item_quantity.values
```

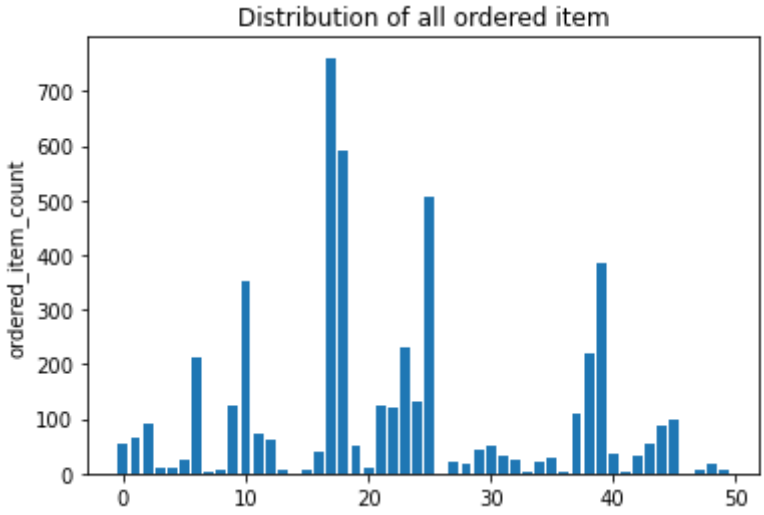
Out[29]:
```
array([ 55,  66,  91,  12,  10,  25, 211,   4,   6, 126, 351,  71,  60,
         8,   1,   6,  40, 761, 591,  50,   9, 123, 120, 230, 130, 506,
         1,  23,  18,  45,  50,  33,  25,   2,  20,  29,   2, 110, 221,
       386,  36,   4,  31,  56,  87,  97,   1,   6,  18,   8], dtype=int64)
```

In [30]:
```
item_quantity.values.tolist()
```

Out[30]:
```
[55,
 66,
 91,
 12,
 10,
 25,
 211,
 4,
 6,
 126,
 351,
 71,
 60,
 8,
 1,
 6,
 40,
 761,
 591,
 50,
 9,
 123,
 120,
 230,
 130,
 506,
 1,
 23,
 18,
 45,
 50,
 33,
 25,
 2,
 20,
 29,
 2,
 110,
 221,
 386,
 36,
 4,
 31,
 56,
 87,
 97,
 1,
 6,
 18,
 8]
```

In [31]:
```
%matplotlib inline
import numpy as np
import matplotlib.pyplot as plt

item_name_list = item_quantity.index.tolist()
x_pos = np.arange(len(item_name_list))
```

```
# 메뉴 이름이 그래프 들어가면 혼잡해지므로 메뉴 이름을 번호화.
order_cnt = item_quantity.values.tolist()

plt.bar(x_pos, order_cnt, align='center')
plt.ylabel('ordered_item_count')
plt.title('Distribution of all ordered item')

plt.show()
```



In [32]:
```
# len(item_name_list) = 50 이므로 x_pos에 저장
np.arange(len(item_name_list))
```

Out[32]:
```
array([ 0,  1,  2,  3,  4,  5,  6,  7,  8,  9, 10, 11, 12, 13, 14, 15, 16,
       17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33,
       34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49])
```

In [33]:
```
type(item_quantity.values)
```

Out[33]:
```
numpy.ndarray
```

In [34]:
```
type(item_quantity.values.tolist())
```

Out[34]:
```
list
```

Problem

In [35]:
```
# values_counts() : item_name 별로 몇 개 있는지  vs  unique() : item_name 이 몇 개 있는
```

In [36]:
```
print(chipo['item_name'].value_counts()[:10])
```

```
Chicken Bowl            726
Chicken Burrito         553
Chips and Guacamole     479
Steak Burrito           368
Canned Soft Drink       301
Chips                   211
Steak Bowl              211
Bottled Water           162
Chicken Soft Tacos      115
Chicken Salad Bowl      110
Name: item_name, dtype: int64
```

In [37]:
```
print(type(chipo['item_name'].value_counts()))
```

```
<class 'pandas.core.series.Series'>
```

In [38]:
```
print(chipo['item_name'].unique()[:10])
```

```
['Chips and Fresh Tomato Salsa' 'Izze' 'Nantucket Nectar'
 'Chips and Tomatillo-Green Chili Salsa' 'Chicken Bowl' 'Side of Chips'
 'Steak Burrito' 'Steak Soft Tacos' 'Chips and Guacamole'
 'Chicken Crispy Tacos']
```

In [39]:
```
print(type(chipo['item_name'].unique()))
```

```
<class 'numpy.ndarray'>
```

전처리

apply(함수), lambda(명령어) 이용

In [40]:
```
print(chipo.info())
print('-----------------------------------')
chipo['item_price'].head()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4622 entries, 0 to 4621
Data columns (total 5 columns):
 #   Column              Non-Null Count  Dtype 
---  ------              --------------  ----- 
 0   order_id            4622 non-null   object
 1   quantity            4622 non-null   int64 
 2   item_name           4622 non-null   object
 3   choice_description  3376 non-null   object
 4   item_price          4622 non-null   object
dtypes: int64(1), object(4)
memory usage: 180.7+ KB
None
-----------------------------------
```

Out[40]:
```
0     $2.39
1     $3.39
2     $3.39
3     $2.39
4    $16.98
Name: item_price, dtype: object
```

In [41]:
```
# 열 단위 데이터에 apply 함수로 전처리 적용
chipo['item_price'] = chipo['item_price'].apply(lambda x : float(x[1:]))
chipo.describe()
```

Out[41]:

	quantity	item_price
count	4622.000000	4622.000000
mean	1.075725	7.464336
std	0.410186	4.245557
min	1.000000	1.090000
25%	1.000000	3.390000
50%	1.000000	8.750000
75%	1.000000	9.250000

	quantity	item_price
max	15.000000	44.250000

In [42]:
```
# apply : chipo['item_price']에 대해 설정
```

In [43]:
```
# float(x[1:]) 의 설명
```

In [44]:
```
'$16.98'[0:]
```

Out[44]: '$16.98'

In [45]:
```
'$16.98'[1:]
```

Out[45]: '16.98'

In [46]:
```
float('$16.98'[1:])
```

Out[46]: 16.98

In [47]:
```
chipo['item_price'].head()
```

Out[47]:
```
0     2.39
1     3.39
2     3.39
3     2.39
4    16.98
Name: item_price, dtype: float64
```

탐색적 분석

In [48]:
```
# 주문당 평균 계산 금액
chipo.groupby('order_id')['item_price'].sum().mean()
```

Out[48]: 18.811428571428717

In [49]:
```
chipo['order_id'] = chipo['order_id'].astype(int)
```

In [50]:
```
chipo.groupby('order_id')['item_price'].mean()
```

Out[50]:
```
order_id
1         2.890000
2        16.980000
3         6.335000
4        10.500000
5         6.850000
           ...
1830     11.500000
1831      4.300000
1832      6.600000
1833     11.750000
```

```
1834     9.583333
Name: item_price, Length: 1834, dtype: float64
```

In [51]:
```
chipo.groupby('order_id')['item_price'].sum()
```

Out[51]:
```
order_id
1       11.56
2       16.98
3       12.67
4       21.00
5       13.70
        ...
1830    23.00
1831    12.90
1832    13.20
1833    23.50
1834    28.75
Name: item_price, Length: 1834, dtype: float64
```

In [52]:
```
chipo.groupby('order_id')['item_price'].sum().describe()[:10]
```

Out[52]:
```
count    1834.000000
mean       18.811429
std        11.652512
min        10.080000
25%        12.572500
50%        16.200000
75%        21.960000
max       205.250000
Name: item_price, dtype: float64
```

In [53]:
```
#한 주문에 10달러 이상 사용한 id 출력
chipo_orderid_group = chipo.groupby('order_id').sum()
results = chipo_orderid_group[chipo_orderid_group.item_price >= 10]

print(results[:10])
print(results.index.values)
```

```
          quantity  item_price
order_id
1                4       11.56
2                2       16.98
3                2       12.67
4                2       21.00
5                2       13.70
6                2       17.50
7                2       15.70
8                2       10.88
9                3       10.67
10               2       13.20
[   1    2    3 ... 1832 1833 1834]
```

In [54]:
```
# 아이템의 가격 계산
# 하나를 주문한 경우만 뽑아보기
chipo_one_item = chipo[chipo.quantity == 1]
# 토핑에 따라서 가격이 다르다.
# 토핑이 올라가지 않은 데이터가 없다면, 어떤 토핑이 올라간 것이 최저가인 게 논리적 비
price_per_item = chipo_one_item.groupby('item_name').min()
price_per_item.sort_values(by="item_price", ascending = False)[:10]
```

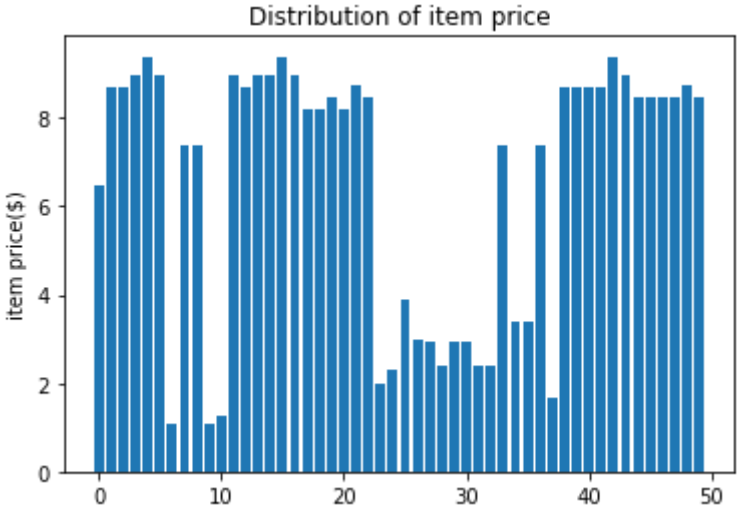
Out[54]:

item_name	order_id	quantity	choice_description	item_price

item_name	order_id	quantity	choice_description	item_price
Steak Salad Bowl	250	1	[Fresh Tomato Salsa, Lettuce]	9.39
Barbacoa Salad Bowl	501	1	[Fresh Tomato Salsa, Guacamole]	9.39
Carnitas Salad Bowl	468	1	[Fresh Tomato Salsa, [Rice, Black Beans, Chees...	9.39
Carnitas Soft Tacos	103	1	[Fresh Tomato Salsa (Mild), [Black Beans, Rice...	8.99
Carnitas Crispy Tacos	230	1	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	8.99
Steak Soft Tacos	4	1	[Fresh Tomato Salsa (Mild), [Cheese, Sour Cream]]	8.99
Carnitas Salad	1500	1	[[Fresh Tomato Salsa (Mild), Roasted Chili Cor...	8.99
Carnitas Bowl	17	1	[Fresh Tomato (Mild), [Guacamole, Lettuce, Ric...	8.99
Barbacoa Soft Tacos	26	1	[Fresh Tomato Salsa, [Black Beans, Cheese, Let...	8.99
Barbacoa Crispy Tacos	75	1	[Fresh Tomato Salsa, Guacamole]	8.99

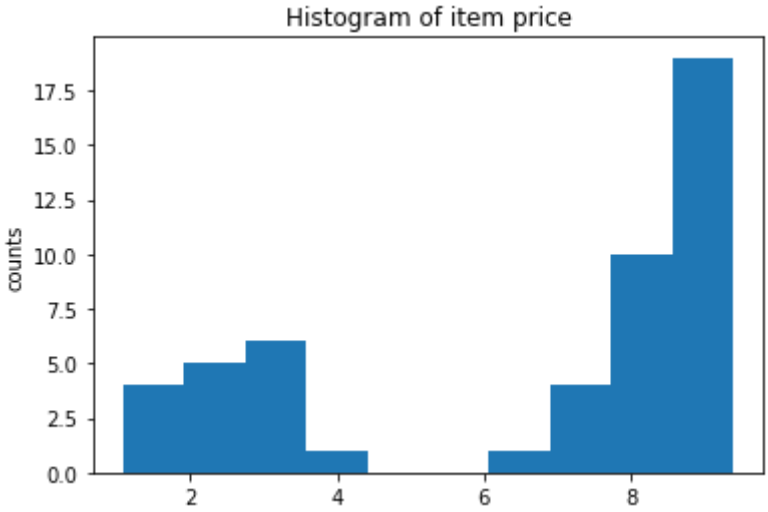
In [55]:

```
# item 가격 분포 그래프
item_name_list = price_per_item.index.tolist()
x_pos = np.arange(len(item_name_list))
item_price = price_per_item['item_price'].tolist()

plt.bar(x_pos, item_price, align='center')
plt.ylabel('item price($)')
plt.title('Distribution of item price')

plt.show()
```



In [56]:

```
# 아이템 가격 히스토그램
plt.hist(item_price)
plt.ylabel('counts')
plt.title('Histogram of item price')

plt.show()
```



In [57]:
```
# 가장 비싼 주문에서 item이 총 몇 개 팔렸는지?
chipo.groupby('order_id').sum().sort_values(by='item_price', ascending=False)[:5]
```

Out[57]:

order_id	quantity	item_price
926	23	205.25
1443	35	160.74
1483	14	139.00
691	11	118.25
1786	20	114.30

In [58]:
```
chipo_salad = chipo[chipo['item_name'] == "Veggie Salad Bowl"]
print(len(chipo_salad))
chipo_salad.head(5)
```

18

Out[58]:

	order_id	quantity	item_name	choice_description	item_price
186	83	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
295	128	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	11.25
455	195	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
496	207	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Rice, Lettuce, Guacamole...	11.25
960	394	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	8.75

In [59]:
```
# “Veggie Salad Bowl” 이 몇 번 주문되었는지를 계산
chipo_salad = chipo[chipo['item_name'] == "Veggie Salad Bowl"]
chipo_salad = chipo_salad.drop_duplicates(['item_name', 'order_id'])
# 한 주문 내에서 중복 집계된 item_name을 제거합니다.(중복된 게 없지만!)
```

```
print(len(chipo_salad))
chipo_salad.head(5)
```

18

Out[59]:

	order_id	quantity	item_name	choice_description	item_price
186	83	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
295	128	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	11.25
455	195	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Rice,...	11.25
496	207	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Rice, Lettuce, Guacamole...	11.25
960	394	1	Veggie Salad Bowl	[Fresh Tomato Salsa, [Fajita Vegetables, Lettu...	8.75

In [60]:

```
# “Chicken Bowl” 을 2개 이상 주문한 주문 횟수를 구합니다.
chipo_chicken = chipo[chipo['item_name'] == "Chicken Bowl"]
chipo_chicken_result = chipo_chicken[chipo_chicken['quantity'] >= 2]
print(chipo_chicken_result.shape[0])
```

33

In [61]:

```
# “Chicken Bowl” 을 2개 이상 주문한 고객들의 "Chicken Bowl" 메뉴의 총 주문 수량을 구
chipo_chicken = chipo[chipo['item_name'] == "Chicken Bowl"]
chipo_chicken_ordersum = chipo_chicken.groupby('order_id').sum()['quantity']
chipo_chicken_result = chipo_chicken_ordersum[chipo_chicken_ordersum >= 2]

print(len(chipo_chicken_result))
chipo_chicken_result.head(5)
```

114

Out[61]:

```
order_id
2      2
34     2
70     2
93     2
124    3
Name: quantity, dtype: int64
```

In [ ]: