Ben Deatsman

Classwork 4b

In [21]:
```
from completejourney_py import get_data
```

In [37]:
```
cj_data = get_data()
cj_data.keys()
```

Out[37]:
```
dict_keys(['campaign_descriptions', 'coupons', 'promotions', 'campaigns', 'demographics', 'transactions', 'coupon_redemptions', 'products'])
```

In [43]:
```
import pandas as pd
products = cj_data['products']
transactions = cj_data['transactions']

transactions.columns.intersection(products.columns)
```

Out[43]:
```
Index(['product_id'], dtype='object')
```

In [50]:
```
transactions.merge(products, how='left', on='product_id')
```

Out[50]:

	household_id	store_id	basket_id	product_id	quantity	sales_value	retail_disc	coupon_disc	coupon_match_disc	week	transaction_timestamp	manufactu
0	900	330	31198570044	1095275	1	0.50	0.00	0.0	0.0	1	2017-01-01 11:53:26	
1	900	330	31198570047	9878513	1	0.99	0.10	0.0	0.0	1	2017-01-01 12:10:28	
2	1228	406	31198655051	1041453	1	1.43	0.15	0.0	0.0	1	2017-01-01 12:26:30	
3	906	319	31198705046	1020156	1	1.50	0.29	0.0	0.0	1	2017-01-01 12:30:27	
4	906	319	31198705046	1053875	2	2.78	0.80	0.0	0.0	1	2017-01-01 12:30:27	
...	...	...	...	...	...	...	...	...	...	...	...	
1469302	679	447	41453103606	14025548	1	0.79	0.20	0.0	0.0	53	2018-01-01 03:50:03	
1469303	2070	311	41453083334	909894	1	1.73	0.17	0.0	0.0	53	2018-01-01 04:01:20	
1469304	2070	311	41453083334	933067	2	5.00	2.98	0.0	0.0	53	2018-01-01 04:01:20	
1469305	2070	311	41453083334	1029743	1	2.60	0.29	0.0	0.0	53	2018-01-01 04:01:20	
1469306	2070	311	41453083334	1061220	1	1.19	0.13	0.0	0.0	53	2018-01-01 04:01:20	

1469307 rows × 17 columns

In [64]:
```
products['department'].unique()
```

Out[64]:
```
array(['GROCERY', 'MISCELLANEOUS', 'PASTRY', 'DRUG GM', 'MEAT-PCKGD',
       'SEAFOOD-PCKGD', 'PRODUCE', 'NUTRITION', 'DELI', 'COSMETICS',
       'MEAT', 'FLORAL', 'TRAVEL & LEISURE', 'SEAFOOD', 'SALAD BAR',
       'FUEL', 'ELECT &PLUMBING', 'FROZEN GROCERY', 'COUPON', 'SPIRITS',
       'GARDEN CENTER', 'TOYS', 'CHARITABLE CONT', 'RESTAURANT',
       'PROD-WHS SALES', 'CHEF SHOPPE', 'GM MERCH EXP', 'AUTOMOTIVE',
       'PHOTO & VIDEO', 'CNTRL/STORE SUP', 'HOUSEWARES', 'POSTAL CENTER'],
      dtype=object)
```

In [78]:
```
(
 transactions
 .merge(products, how='left', on='product_id')
 .query("department == 'MEAT'")
 .groupby('product_category', as_index=False)
 .agg({'sales_value': 'sum'})
 .sort_values(by='sales_value', ascending=False)
)
```

Out[78]:

	product_category	sales_value
1	BEEF	176614.54
2	CHICKEN	52703.51
10	PORK	50809.31
12	SMOKED MEATS	15324.22
13	TURKEY	11128.95
4	EXOTIC GAME/FOWL	860.42
5	LAMB	829.27
14	VEAL	167.13
8	MEAT SUPPLIES	57.03
7	MEAT - MISC	39.66
11	RW FRESH PROCESSED MEAT	30.84
3	COUPON	7.00
6	LUNCHMEAT	2.20
9	MISCELLANEOUS	0.95
0	BACON	0.30

In [ ]:
```

```