

Problem :

GODIGT Bank also has a set of customers who were given credit cards based on risk policy and customer category class but due to huge competition in the credit card market, the bank is observing high attrition in credit card spending. The bank makes money only if customers spend more on credit cards. Given the attrition, the Bank wants to revisit its credit card policy and make sure that the card given to the customer is the right credit card. The bank will make a profit only through the customers that show higher intent towards a recommended credit card. (Higher intent means consumers would want to use the card and hence not be attrite.)

QUESTION: Analyse the dataset and list down the top 5 important variables, along with the business justifications.

SOLUTION: Dataset: godigt_cc_data

GODIGT Bank is a mid-sized private bank that deals in all kinds of banking products, such as savings accounts, current accounts, investment products, etc. among other offerings.

- The dataset contains 8448 rows and 28 columns.
- Dataset contains columns having data type int, object and datetime as follows:

S.NO.	COLUMN NAME	DATA TYPE
1	userid	int64
2	card_source_date	datetime64
3	widget_products	int64
4	engagement_products	int64
5	active_30	int64
6	active_60	int64
7	active_90	int64
8	Issuer	object
9	card_type	object
10	hotlist_flag	object
11	avg_spends_13m	object

After taking a look on the dataset:

- The columns having datatype “object” has unique values:

S.NO.	UNIQUE VALUES	COLUMN NAME	DATA TYPE
1	Visa, Mastercard, Amex	Issuer	object
2	edge, prosperity, rewards, Indianoil ,et c.	card_type	object
3	B, A, C, E, D	high_networth	object
4	N, Y	hotlist_flag	object
5	N, Y	other_bank_cc_holding	object
6	T,R	Transactor_revolver	object
7	Self Employed, 0, Student, Salaried, Retired, Housewife	Occupation_at_source	object

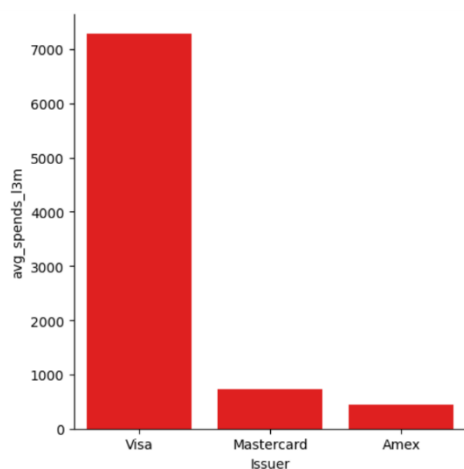
- One column has missing values:

COLUMN NAME	NO. OF MISSING VALUE
Transactor_revolver	38

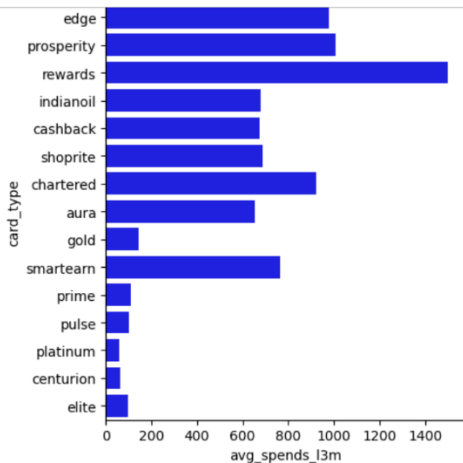
The top 5 important variables, along with the business justifications:

1. The credit card name:

```
! : sns.catplot(df,x=df.Issuer,y=df.avg_spends_l3m ,kind='bar',errorbar=None,estimator='count', palette=['red'])
```



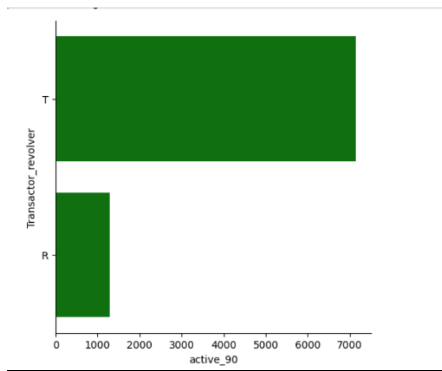
```
: sns.catplot(df,y=df.card_type,x=df.avg_spends_l3m ,kind='bar',errorbar=None,estimator='count', palette=['blue'])
```



As we can see from the above charts that only four-five cards are in high demand. Therefore the companies should focus only on such cards namely rewards, edge etc to increase its sales. Also the companies can launch some offers on these cards to attract customers which uses other card.

2. Transactor or Revolver:

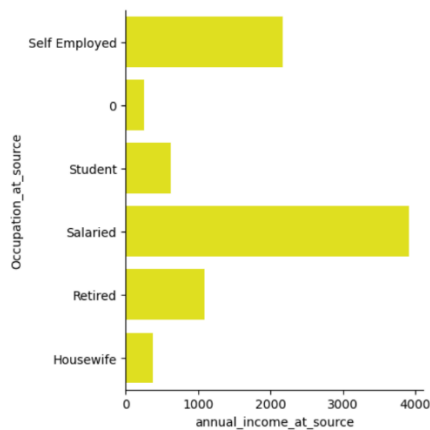
```
sns.catplot(df,x=df.active_90,y=df.Transactor_revolver ,kind='bar',errorbar=None,estimator='count', palette=['blue'])
```



The above graph indicates the no. of consumers who constantly forward their balance from the last three months. The companies should definitely take a keen look towards this as it is evenly important for the business. The companies can use cold-calling otherwise can imposed penalties.

3. Occupation:

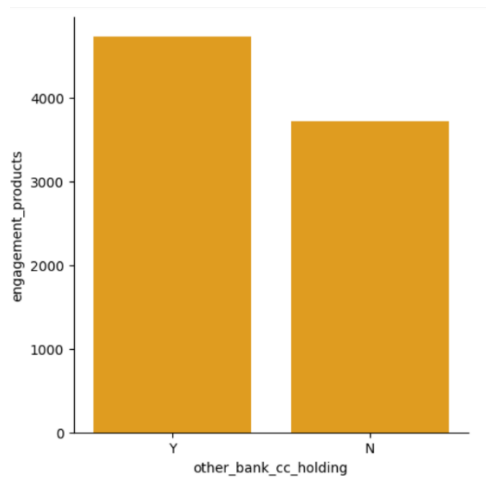
```
sns.catplot(df,y=df.Occupation_at_source,x=df.annual_income_at_source ,kind='bar',errorbar=None,estimator='count', palette=['yellow'])
```



It is clear from the bar graph that salaried class people generate highest income followed by self-employed and retired persons. Companies should decrease the ROI for students & housewives as they have less or no income to increase its revenue also they can offer some rewards points or coins to their highest revenue generator's so that they might not lose interest.

4. Hold other bank credit card

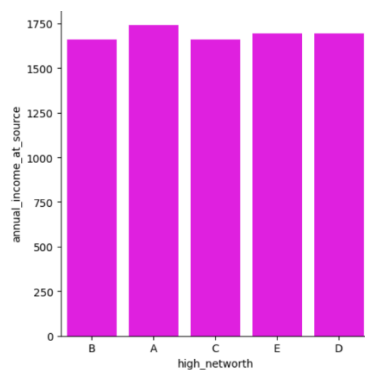
```
•[55]: sns.catplot(df,y=df.engagement_products,x=df.other_bank_cc_holding,kind='bar',errorbar=None,estimator='count', palette=['yellow'])
```



Companies should analyses whether the customer has other bank credit card before selling their investment/loan product (FD, RD, Personal loan, auto loan etc) .If the customer has other bank card he/ she should be given less products and also on slightly higher interest as there is possibility of higher risk.

5. Annual income by Customer category:

```
: sns.catplot(df,y=df.annual_income_at_source ,x=df.high_networth ,kind='bar',errorbar=None,estimator='count', palette=['green'])
```



The graph shows that customers which comes under the category of “A” generates highest income for the bank followed by “D” and “C”. So companies should focus on these customers.