# Absa Customer Income Prediction Challenge

Can you predict a customer's income based on their transaction history?

Prize

**$5 000 USD**

Time

**3 months to go**

Participants

**14 active · 196 enrolled**

Helping

**South Africa**

Good for beginners Prediction Financial Services



# About

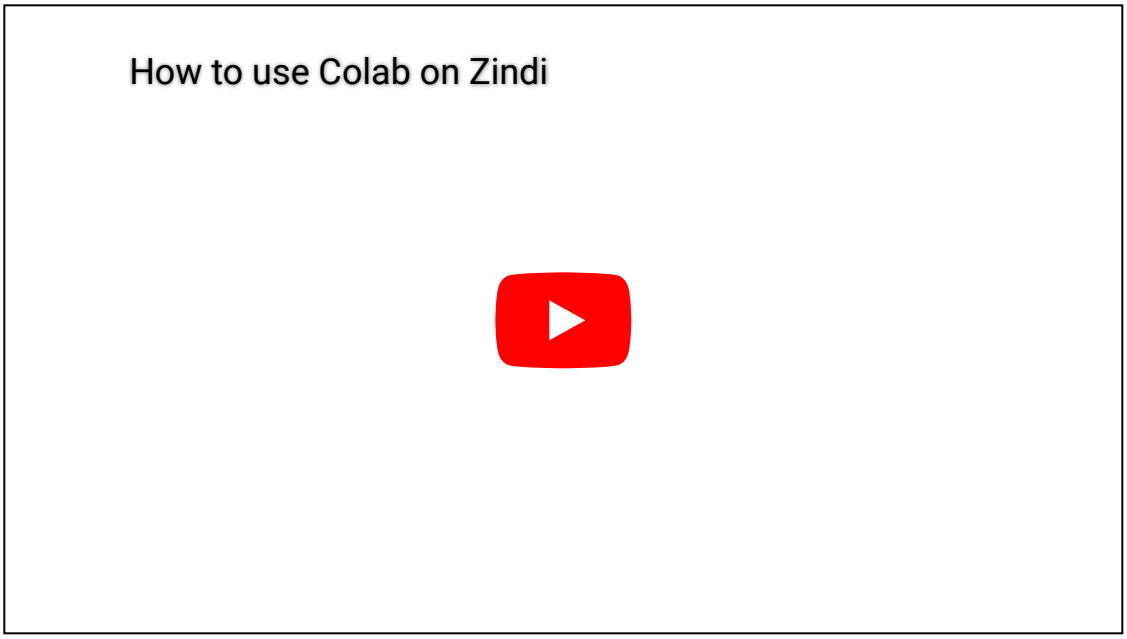
The transaction data is over a period of 14 months, running from the start of July 2021 to the end of August 2022. There are 5144 customers and 46926 unique accounts, with some customers having more than one account. The number of transactions per customer range from 1 to just over 2 000 for the 14-month period.

The train set contains 3600 customers along with their declared net income and the test data contains the remaining 1544 customers with the declared income excluded.

Along with the train and test files is the transaction history for each customer which details the

data files containing demographic information on each customer as well as files describing some of the categorical variables related to the customers.

### How to use Colab on Zindi




### How to mount a drive on Colab


The objective of this challenge is to create a machine learning solution to determine a customer's income based on their transaction history over 14 months. Your model can make use of all the data provided with the target being the declared net income from the train file.

- *employment_status.csv* - contains descriptions of the various categories of customer employment status
- *income_group.csv* - describes how customers are grouped based on their income
- *Test.csv* - contains customer identifiers as well as record dates for customers in the test set
- *SampleSubmission.csv* - a sample of the test file with zero predictions on the declared net income
- *Train.csv* - contains customer identifiers as well as their declared net income(the target) used to train your model
- *customer.csv* - contains demographic information on the customers that can be used as predictors in your model
- *transactions.csv* - customer transaction history for the past 14 months

# Files

Contains customer identifiers as well as record dates for customers in the test set

Test.csv
211.1 KB

Contains customer identifiers as well as their declared net income(the target) used to train your model

Train.csv
522.3 KB

Describes how customers are grouped based on their income

income_group.csv
463 B

Contains demographic information on the customers that can be used as predictors in your model

customer.csv
763.8 KB

Contains descriptions of the various categories of customer employment status

employment_status.csv
83 B

This shows the submission format for this competition, with the ‘CUSTOMER_IDENTIFIER’ column mirroring that of Test.csv and the ‘DECLARED_NET_INCOME’ column containing your predictions. The order of the rows does not matter, but the names of the ‘CUSTOMER_IDENTIFIER’ must be correct.

SampleSubmission.csv
202.1 KB

Customer transaction history for the past 14 months

transactions.csv
427.3 MB

## Events

Competitions

## Learn

Learning Center

Discussions

## Opportunities

Jobs Board

## Contacts

zindi@zindi.africa

## Info

About Us

FAQs

Our Partners

## Business

Host competition