WorldCover

Business Intelligence

MINI PROJECT

Please return your findings by end of day **Sunday, July 28** (EDT).

Background

This mini project is based on a real-life use case using a sample of WorldCovers transactional database.

A particular business challenge for our Customer Success team is to increase customer retention season after season and to track and limit customer churn (i.e. limit the number of farmers who decide not to renew their drought insurance contract from one season to the next).

Your role as our BI Lead is to:

1) Define key metrics that can inform our CS Team about **customer retention trends** (here we think of a "retained" customer as a customer who purchased drought insurance in multiple cropping seasons)

2) Present this information in a **synthetic and visual manner** (you might be asked to present some of these findings orally during our final interview)

Some sample questions that are of particular interest to our team:

1. Do retention patterns vary by geographic zone?

2. Are retention rates and insurance premium amounts influenced by external factors, such as customers experiencing drought (and receiving a payout) in a prior cropping season?

3. Do other socio-demographic factors have an influence on customer retention patterns?

Instructions

To investigate these questions we provide you with table dumps in CSV format, and with a data schema that documents table fields and relationships. You are by no means limited to these datasets and you can use any external sources you think can help bring additional insights about our customers' behavior.

You have until **Thursday midnight (EDT)** to complete this project. You can dedicate as much or as little time as you wish/need (but really we don't expect you to spend any longer than 3 to 4hrs). Reach out to melanie.bacou@worldcovr.com or Skype (mbacou) if you have any questions.

- We expect the main deliverable to be an **interactive dashboard** (or screenshot) using any tool or platform of your choice. A static document, static presentation, or spreadsheet model are all acceptable, as long as your document your process.
- Document and explain problems you encounter along the way and choices you make, what pitfalls, if any, you see with our current data model. Also tell us what other questions around customer retention you would like to investigate given more time and more data.
- Finally, provide any **queries/code files** that you used to generate all artifacts.
- Your solution to this project should live in a **publicly accessible Git repository**.

Don't worry! This is an open-ended exercise and is meant to test whether you can hack together metrics and a visual story on a small part of our production data.

Good luck!

Definitions

`season`	Agricultural cropping season (last about 6 months). Some regions have 2 seasons in a year, a major or long rainy season from March to August and a minor or short rainy season from July/August to November/December
`premium`	Cost of drought insurance (amount paid by a customer over cash or USSD)
`payout`	Claim payment made by the insurer to the policyholder when drought events occur. In Ghana payouts are made to entire communities in one lump sum transaction.
`retention rate`	Percentage of prior season customers who purchase insurance again in a subsequent season
`churn rate`	Percentage of prior season customers who do not purchase insurance again in a subsequent season
`USSD`	Unstructured Supplementary Service Data, also called "Quick Codes" or "Feature codes", is a text-based communications protocol used by GSM cellular phones to communicate with the mobile network operator's computers. USSD can be used for prepaid callback service, mobile-money services, location-based content services, and menu-based information services.

Data Schema

5 tables in CSV format are included:

communities

`community_id`	Unique identifier
`community_name`	Name of the community
`country`	Country name
`region`	Top level subdivision in the country
`district`	Second level subdivision in the country (contained in a region)
`latitude`	Community GPS latitude in decimal degree
`longitude`	Community GPS latitude in decimal degree

customers

`customer_id`	Unique identifier
`community_id`	Foreign key to the `communities` table referencing where the customer is located
`registration_date`	The date that the customer first signed up with WorldCover
`created_at`	Time customer record was created
`gender`	M (male) or F (female)
`has_phone`	Customer has a mobile phone
`farm_size`	Customer farm size in acres
`literacy`	Customer literacy level
`ussd_created`	Customer was first registered through USSD (mobile phone)

customer_policies

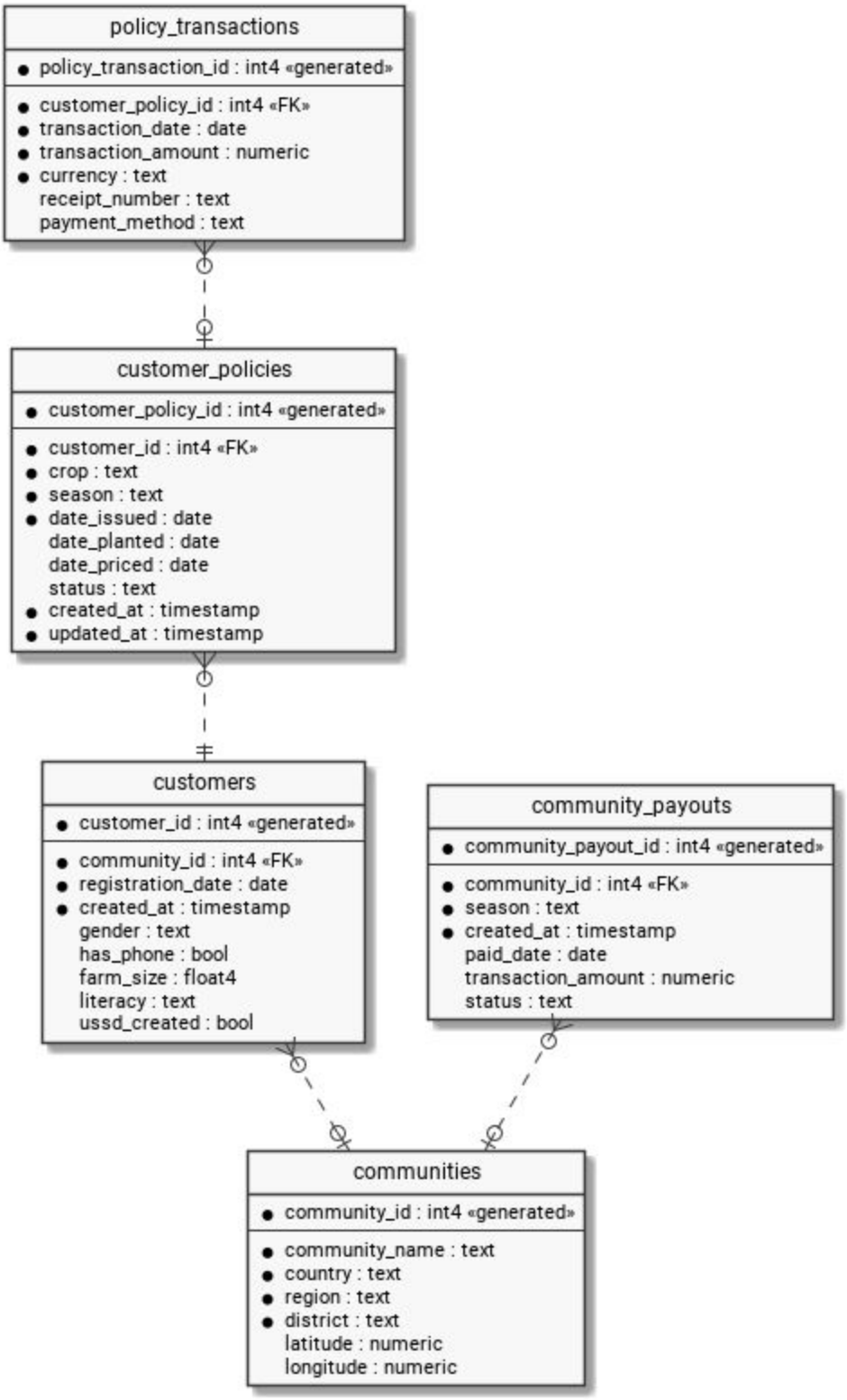
`customer_policy_id`	Unique identifier
`customer_id`	Foreign key linking to the `customers` table referencing which customer purchased this policy
`crop`	The crop that is insured by this policy
`season`	The cropping season year that was covered by the policy. In certain regions there are 2 cropping seasons in a year (major & minor)
`date_issued`	The date that the customer first purchased this policy
`date_planted`	The date that the farmer planted the crop. This determines the trigger start date for the policy
`date_priced`	Date when the policy was priced
`status`	Current policy status; one of `pending`, `planted`, `priced`, `active`, `triggered`, `expired`, `payout due`, `payout initiated`, `paid out`, `dispute`, `refunded`

policy_transactions

policy_transaction_id	Unique identifier
customer_policy_id	Foreign key linking to the customer_policies table referencing which policy this payment was associated with
transaction_date	The date that the customer made the payment towards the policy
transaction_amount	The amount of the payment in local currency
currency	The currency used for the payment
receipt_number	The receipt for the transaction
payment_method	Method of payment

community_payouts

community_payout_id	Unique identifier
community_id	Foreign key linking to the communities table
season	The cropping season year that was covered by the policy. In certain regions there are 2 cropping seasons in a year (major & minor)
created_at	Time when payout record was first created
paid_date	The date that the payout was approved and sent to the customers
transaction_amount	The amount of the payment in local currency
status	Payment status



WorldCover Data Schema (Ghana, extracts)