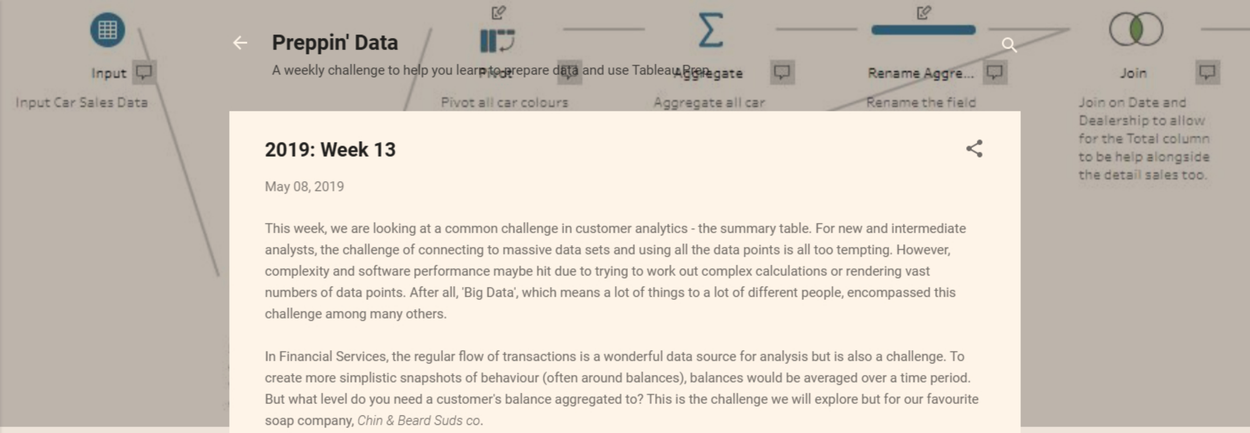


2019: Week 13

May 08, 2019

This week, we are looking at a common challenge in customer analytics - the summary table. For new and intermediate analysts, the challenge of connecting to massive data sets and using all the data points is all too tempting. However, complexity and software performance maybe hit due to trying to work out complex calculations or rendering vast numbers of data points. After all, 'Big Data', which means a lot of things to a lot of different people, encompassed this challenge among many others.

In Financial Services, the regular flow of transactions is a wonderful data source for analysis but is also a challenge. To create more simplistic snapshots of behaviour (often around balances), balances would be averaged over a time period. But what level do you need a customer's balance aggregated to? This is the challenge we will explore but for our favourite soap company, *Chin & Beard Suds co.*

Using three different company's data who buy our products, we are looking at the balance of credit that they hold with us. As any business owner will know, 'cashflow is king' and therefore, we only provide supplies to those who hold a positive balance with us at the start of the month. But are these companies 'gaming' our system? We want to know:

- What is the average weekly, monthly and quarterly balance?
- What is the average weekly, monthly and quarterly transaction value?
- How many days does the customer have a negative balance?
- How many days does the customer exceed their credit limit? (credit limit is a positive number in the input but needs to be made negative as it how much we allow the customer to owe us)

Requirements

Account	Name	Max Credit
1237421	Bubbly McBubbleface	10000
4271819	Bubblicious	30000
12371202	Bubbleburst	5000

Input of Customer Details

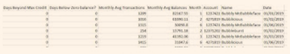
Account	Date	Transaction	Balance
1237421	01/01/2019	578	100000
1237421	02/01/2019	198	99422
1237421	03/01/2019	1806	99224
1237421	04/01/2019	144	97418
1237421	05/01/2019	2240	97274
1237421	06/01/2019	1773	95034

Input of Transactions

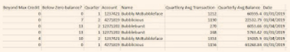
- Input data
- Create an average (mean) for balance to two decimal places and average (mean) for transactions per customer to no decimal places, per time period
- Each row will be a customer per time period
- Date recorded will be the beginning of that time period (ie first day in the week for the weekly table)
- Bring in the customer name
- Determine the number of days a customer's balance is below zero
- Determine the number of days a customer's balance is below their credit limit (ie have gone beyond our allowance).

Output

Weekly



Monthly



Quarterly

- 3 Output files: weekly, monthly, quarterly
- Outputs (excluding headers in counts):
  - Weekly: 81 rows, 8 columns
  - Monthly: 18 rows, 8 columns
  - Quarterly: 6 rows, 8 columns

For comparison, here's our output files. Don't to forget to fill in our participation tracker!

Aggregation clean Join

Popular posts from this blog

2023: Week 1 The Data Source Bank

January 04, 2023

Created by: Carl Allchin Welcome to a New Year of Preppin' Data. These are weekly exercises to help you learn and develop data preparation skills. We publish the challenges on a Wednesday and share a solution the following Tuesday. You can take the challenges whenever you want and we love to see y ...

READ MORE

2023: Week 2 - International Bank Account Numbers

January 11, 2023

Challenge By: Jenny Martin For week 2 of our beginner month, Data Source Bank has a requirement to construct International Bank Account Numbers (IBANs), even for Transactions taking place in the UK. We have all the information in separate fields, we just need to put it altogether in the following order: ...

READ MORE

2021: Week 22 - Answer Smash

June 02, 2021

Challenge By: Jenny Martin Recently, my family and I have become quite invested in the TV quiz show Richard Osman's House of Games . The final round is always a round called Answer Smash. In this round you have a picture and question and you have to "smash" the name of the picture with the answ ...

READ MORE