

MICHIGAN STATE

U N I V E R S I T Y

Alpha Presentation

Excel Data Extractor and Modeler

The Capstone Experience

Team Roosevelt Innovations Knowledge Science

Brendan Bushbaker

Het Patel

Kathryn Nagy

Rithwik Pulichera

Adrian Self

Department of Computer Science and Engineering

Michigan State University

Spring 2024



*From Students...
...to Professionals*

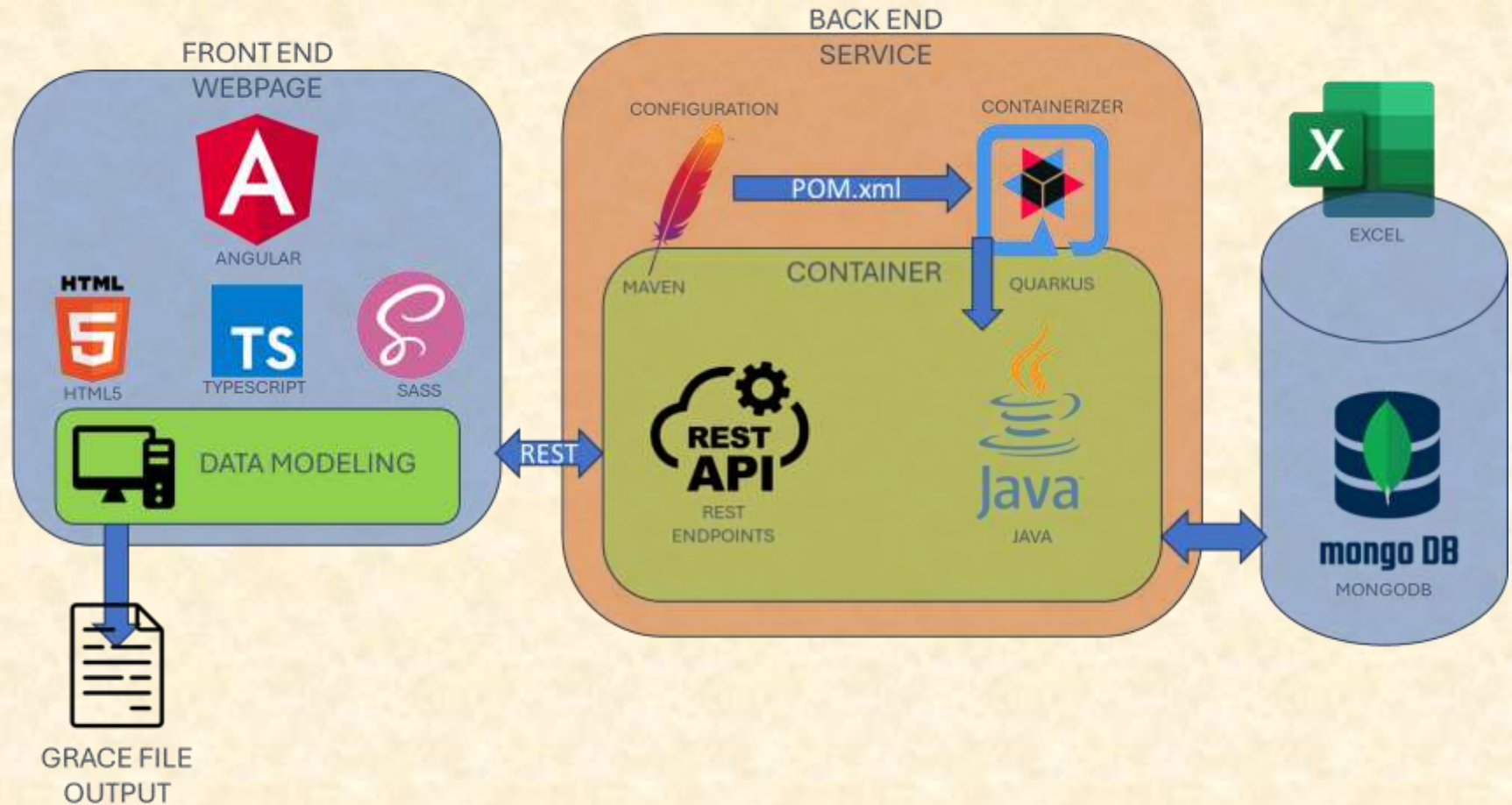
Project Overview

- Provides SaaS for Insurance Analysts with 50+ years of claims experience
- Have saved about 972 million dollars in treatment costs
- Founded as a subsidiary of Delta Dental of Michigan
- Want to provide a better way to process their claims rate calculations
- A better way to extract and model their claims data from Excel



System Architecture

MICROSOFT EXCEL DATA EXTRACTOR/MODELER



Labeling and Measurements

The screenshot displays a web application interface for data analysis. The main area is a data table with columns labeled A through P. The table is organized into sections for different states: Michigan (rows 3-11) and Ohio (rows 12-20). Each state section includes rows for 'Individual' and 'Spouse' plan types, with 'Dependent' rows also present. The columns represent time periods: 2021 (Q1, Q2, Q3) and 2022 (Q1, Q2, Q3). A 'CustomersPerQuarter2020-2021' column is also visible. The interface includes a 'Social Data extractor' browser window at the top, a 'Roosevelt' logo with the tagline 'simple. seamless. smart.', and a sidebar on the right with 'Added Labels' and 'Suggested Labels' sections. The 'Added Labels' section contains labels like 'Temp - Range: D3:L11', 'States - Range: A2:A29', and 'TestingLabel - Range: L6:6'. The 'Suggested Labels' section contains labels like 'Premier - Range: B3:J3', 'Indiana - Range: D21:L29', 'Ohio - Range: D12:L20', 'PPO - Range: B2:J5', 'Michigan - Range: D31:L11', and 'OCN - Range: B4:J4'. An 'OPTIONS' button is located at the bottom center of the table.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1				2021			2022			2023						000	
2				Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3				Name	
3	Michigan	Individual	PPO	2.31	5.92	5.21	1.20	7.35	8.90	4.52	5.80	7.80				State	
4			Premier	3.92	6.74	3.91	6.10	6.51	1.73	3.73	1.73	3.71				Sub Type	
5			OCN	4.17	4.95	5.17	1.47	6.36	6.41	3.43	2.89	5.17				Net Type	
6		Spouse	BPD	7.64	1.11	5.48	1.88	3.34	7.00	3.45	5.11	5.48				Age	
7			Premier	2.81	5.78	7.98	2.21	2.58	4.85	6.78	5.78	7.88					
8			OCN	2.01	5.73	1.68	5.83	3.74	5.88	2.81	5.73	1.68					
9		Dependent	PPO	1.38	4.30	4.57	3.51	5.16	8.12	2.01	4.30	4.57					
10			Premier	7.48	4.45	6.38	4.81	6.84	1.81	1.39	4.45	4.39			CustomersPerQuarter2020-2021		
11			OCN	6.38	2.80	3.51	1.68	1.38	4.80	7.48	2.80	3.51			2020	400	
12		Ohio	Individual	BPD	7.13	6.11	4.97	6.16	1.42	4.00	5.38	6.11	4.97			2020	300
13				Premier	1.80	1.27	4.01	3.06	4.80	4.02	7.13	1.27	4.01			2020	350
14	OCN			2.58	4.63	2.63	8.10	1.26	1.40	1.80	4.63	3.03			2020	400	
15	Spouse		PPO	1.25	7.44	1.91	6.45	4.17	3.00	3.58	7.44	1.91			2021	840	
16			Premier	6.44	5.71	4.17	4.75	4.69	6.51	1.25	5.71	4.17			2021	700	
17			OCN	7.42	4.02	1.08	2.53	6.99	1.89	6.44	4.02	1.08			2021	760	
18	Dependent		PPO	2.97	2.75	4.26	1.88	2.49	6.25	7.42	2.75	4.26			2022	850	
19			Premier	5.66	1.17	7.24	2.16	4.11	5.08	3.97	1.17	7.24					
20			OCN	5.75	4.96	3.18	2.67	3.84	5.72	5.66	4.96	3.18					
21				PPO	1.44	2.76	7.18	4.15	2.55	6.37	1.76	7.18					



Formula Analysis

The screenshot shows an Excel spreadsheet with the following data:

	A	B	C	D	E	F	G
1		Apple	Orange	Lemon	Totals	Products	Combination
2	Tasty	54	58	54	166	4344	144364260976
3	Sour	26	9	67	102	15678	
4	Plain	10	42	5	57	2109	

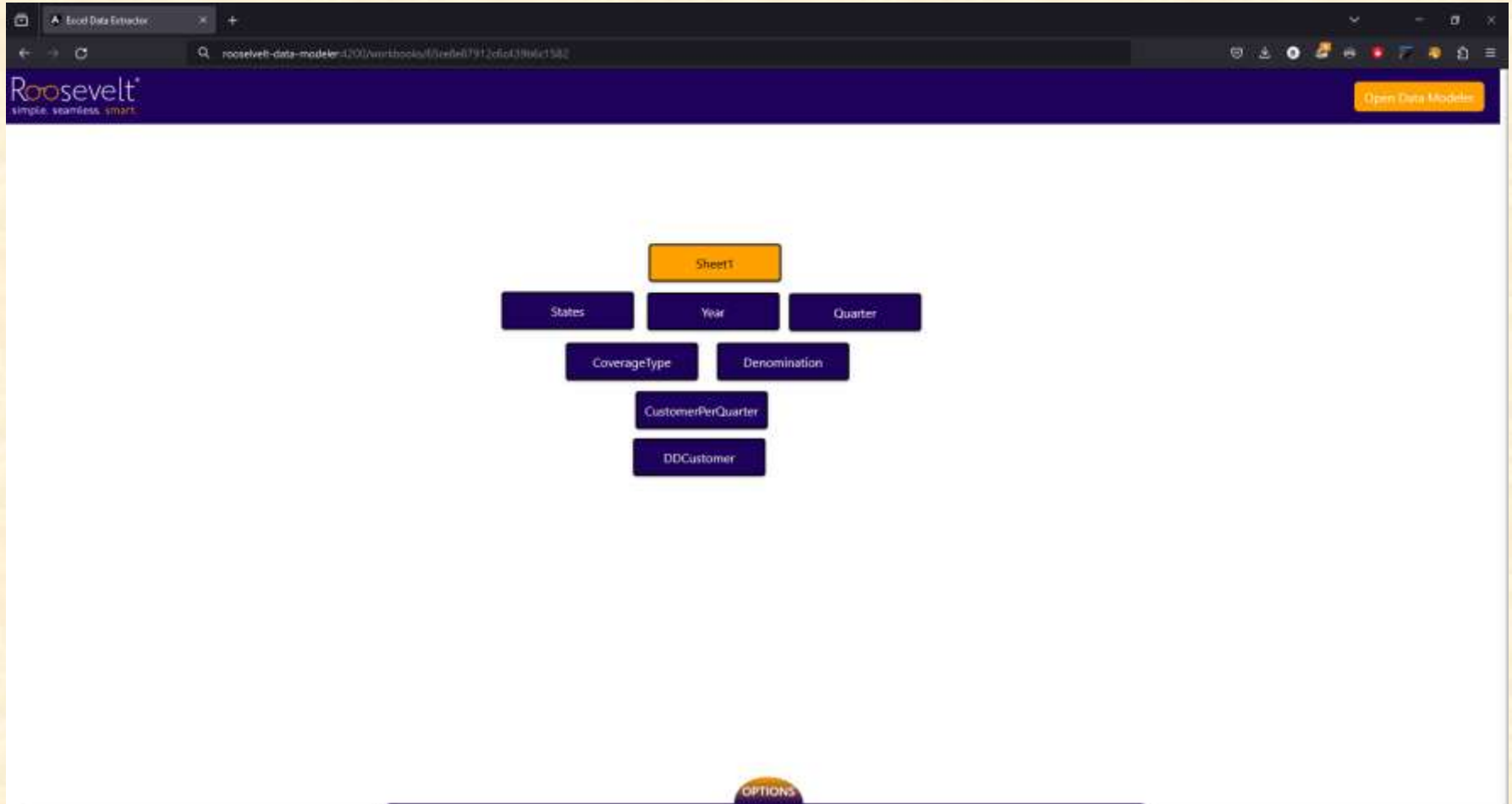
Below the spreadsheet, a formula analysis window is displayed for cell G2. It shows the function signature:

Function Signature:
SUM(D2:E2) + PRODUCT(B2:D4) - PRODUCT(E2:E4)

The corresponding mathematical expression is:

$$\sum_{i=1}^2 x_i + \prod_{i=1}^9 x_i - \prod_{i=1}^3 x_i$$


Data Modeler



Workbook Selection

The screenshot displays the Roosevelt Excel Data Extractor web application. The browser's address bar shows the URL: `roosevelt-data-extractor.4200/workbooks/55a6e679126643866c1582`. The page header includes the Roosevelt logo and the text "Excel Data Extractor". In the top right corner, the user's name "Jacob" is displayed next to a "SIGN OUT" button.

The main content area is titled "Spreadsheets" and contains a table with the following data:

Name	Date Created	Option
Fruits3.xlsx	Feb 18, 2024	Delete Model
Fruits3.xlsx	Feb 18, 2024	Delete Model
DeltaDentalGraceTester.xlsx	Feb 15, 2024	Delete Model

Below the table, there is a pagination control showing "Items Per Page" set to 10 and "1 - 3 of 3" items. A yellow "UPLOAD WORKBOOK" button is located at the bottom right of the spreadsheet list area.



What's left to do?

- Finish data modeler.
- Support for additional formulas.
- Upgrade suggested labels.
- Tree view selection.
- UI updates and fixes.
- GRACE file validation.



Questions?

?

?

?

?

?

?

?

?

?

