

Procurement Fraud Analysis

Project Scope:

- Marketing Procurement is the buying of goods and services that support an organization's Marketing investment.
- Procurement fraud is a deliberate deception intended to influence any stage of the procure-to-pay lifecycle in order to make a financial gain or cause a loss.
- Procurement fraud is considered one of the most common frauds perpetrated by officers, vendors operating in emerging markets, and the big losers of this fraud are businesses.
- **Technical stack:**
- We have used Tableau to build the visualizations and the dashboard.
- The granularity of data in different segments can be visualized.

Project Goals:

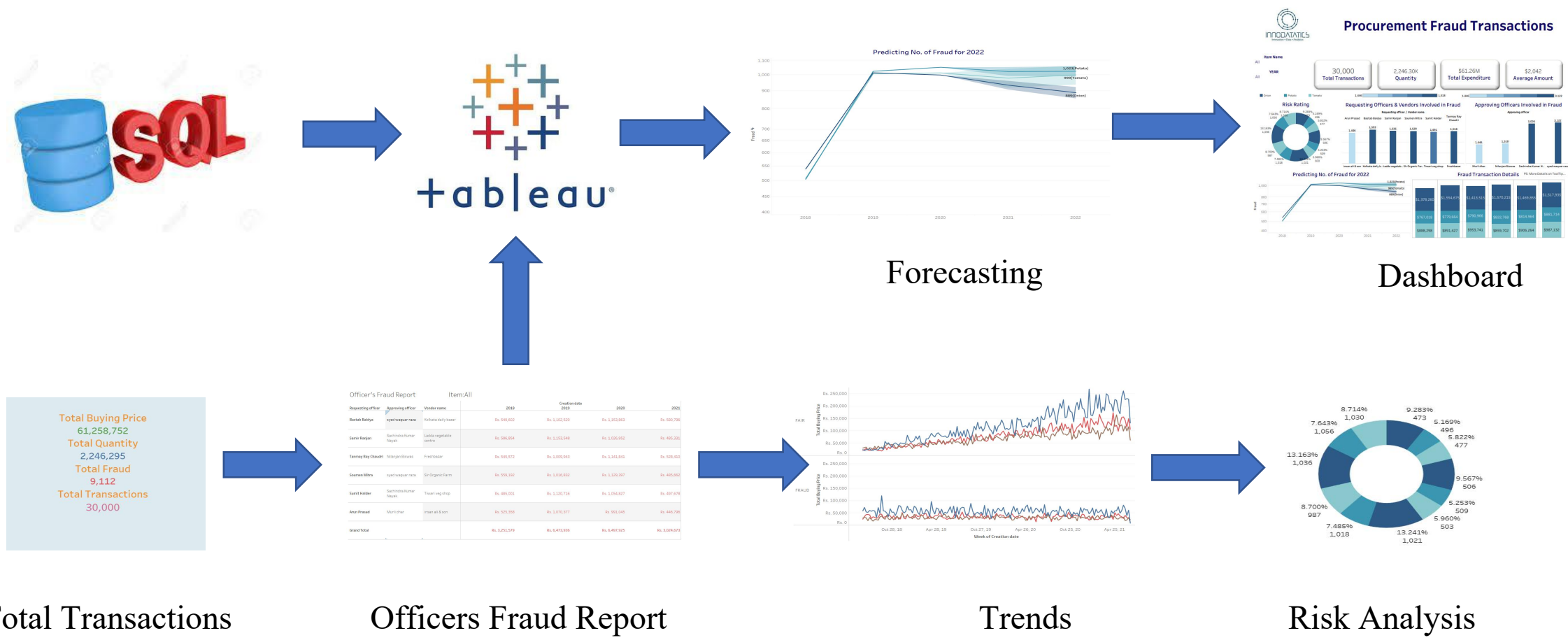


- Analysis of Sales and Profit.
- Detection of Anomalous behavior.
- Detect insider enabled procurement Fraud.
- Fraud Risk Assessment to identify and Mitigate Fraud.



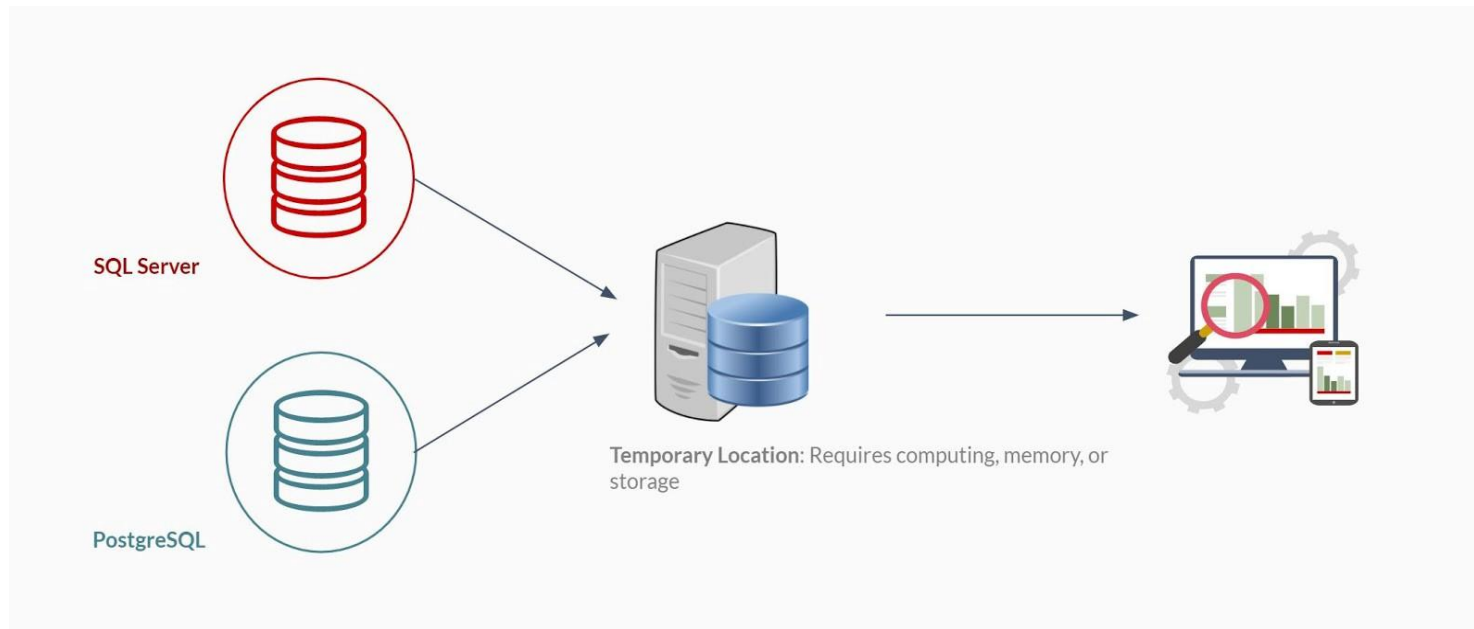
- The constraints of project is to analyze the year where number of fraud happened is more.
- Analyze the Product item sales is high in which year.

Project Architecture:



Data Understanding:

- Determine what data is needed.
- Collect data, if required data is not available.
- Explore data.
- Verify data quality.



Data Preparation:

Data Collection:

- The Data for doing the analytics is taken from the SQL server.
- Dataset shape : observations-30001, features-11.

Data Types:

- The given data set contains 11 variables as follows:
- Creation date : Time series data.
- PO_Id, Requesting officer, Approving officer, PO_item, Item number, Quantity, Unit_max_price, Buying price, Vendor name, Total Buying Price Fraud - perform all these variables of continuous data.



SQL



Tableau Prep

Data Preparation:

- Clean data.
- Generate new attributes required for analysis (also called *Feature Engineering*).
- Integrate data (merge and aggregate).

Original Data:

New Data:

This screenshot shows an Excel spreadsheet titled 'Data - Excel (Product Activation Failed)'. The spreadsheet contains a table with 9 columns: Creation_date, PO_item, Item_number, Quantity, Unit_max_price, Buying_price, Vendor_name, Total Buying Price, and Fraud. The data is organized into 29 rows, with the first row being the header. The 'Fraud' column contains values of 0 for all rows. The 'Total Buying Price' column contains values ranging from 1760 to 3185. The 'Buying_price' column contains values ranging from 14 to 65. The 'Quantity' column contains values ranging from 14 to 96. The 'Unit_max_price' column contains values ranging from 25 to 60. The 'Vendor_name' column contains various vendor names such as 'Garden fresh', 'puja vegetable shop', 'Freshbazar', 'Rupdasi market', 'Green Tokri', 'Abu variety stores', 'Newtown home delivery', 'Once more', 'Kolkata Mart', 'Kolkata daily bazar', 'RK wholesale vegetable', 'Newtown home delivery', 'Abu variety stores', 'Quadri vegetable shop', 'mohidul vegetable shop', 'Halder vegetarian', 'Kolkata daily bazar', 'Ladda vegetable centre', 'Rupdasi market', and 'Tiwaree vee shoo'. The 'PO_item' column contains values such as '1/19/2021 Onion', '2/14/2019 Tomato', '8/13/2019 Onion', '4/30/2021 Tomato', '5/13/2020 Onion', '6/21/2021 Onion', '11/4/2018 Tomato', '9/13/2020 Tomato', '2/3/2019 Onion', '7/14/2019 Tomato', '1/31/2021 Onion', '10/5/2020 Tomato', '9/29/2018 Onion', '10/16/2018 Onion', '2/23/2019 Onion', '1/10/2019 Tomato', '6/28/2021 Onion', '6/6/2021 Potato', '4/14/2019 Potato', '9/12/2020 Tomato', '1/26/2020 Onion', '3/8/2021 Onion', '8/8/2020 Tomato', '8/17/2019 Potato', '3/5/2020 Tomato', '10/31/2019 Tomato', '8/28/2020 Onion', and '4/18/2021 Onion'. The 'Item_number' column contains values ranging from 52 to 73. The 'Creation_date' column contains dates ranging from 1/19/2021 to 4/18/2021. The 'Quantity' column contains values ranging from 14 to 96. The 'Unit_max_price' column contains values ranging from 25 to 60. The 'Buying_price' column contains values ranging from 14 to 65. The 'Vendor_name' column contains various vendor names. The 'Total Buying Price' column contains values ranging from 1760 to 3185. The 'Fraud' column contains values of 0 for all rows. The spreadsheet is displayed in 'Sheet1' with a status bar at the bottom showing 'AVERAGE: 7658.912879 COUNT: 2124 SUM: 8087812'.

Creation_date	PO_item	Item_number	Quantity	Unit_max_price	Buying_price	Vendor_name	Total Buying Price	Fraud
1/19/2021	Onion	GRO001	52	60	58	Garden fresh	3016	0
2/14/2019	Tomato	GRO002	63	14	14	puja vegetable shop	882	0
8/13/2019	Onion	GRO001	63	30	30	Freshbazar	1890	0
4/30/2021	Tomato	GRO002	62	41	41	Rupdasi market	2542	0
5/13/2020	Onion	GRO001	53	45	44	Rupdasi market	2332	0
6/21/2021	Onion	GRO001	70	65	65	Green Tokri	4550	0
11/4/2018	Tomato	GRO002	51	11	10	Abu variety stores	510	0
9/13/2020	Tomato	GRO002	89	32	31	Newtown home delivery	2759	0
2/3/2019	Onion	GRO001	73	20	19	Rupdasi market	1387	0
7/14/2019	Tomato	GRO002	96	20	18	Once more	1728	0
1/31/2021	Onion	GRO001	92	60	56	Kolkata Mart	5152	0
10/5/2020	Tomato	GRO002	65	35	34	Kolkata daily bazar	2210	0
9/29/2018	Onion	GRO001	51	25	13	Rupdasi market	663	0
10/16/2018	Onion	GRO001	83	25	12	RK wholesale vegetable	996	0
2/23/2019	Onion	GRO001	88	20	20	Newtown home delivery	1760	0
1/10/2019	Tomato	GRO002	60	14	12	Abu variety stores	720	0
6/28/2021	Onion	GRO001	88	65	63	Quadri vegetable shop	5544	0
6/6/2021	Potato	GRO003	88	32	31	mohidul vegetable shop	2728	0
4/14/2019	Potato	GRO003	68	16	16	Abu variety stores	1088	0
9/12/2020	Tomato	GRO002	91	32	31	Tiwari veg shop	2821	0
1/26/2020	Onion	GRO001	96	40	37	Green Tokri	3552	0
3/8/2021	Onion	GRO001	65	60	60	Abu variety stores	3900	0
8/8/2020	Tomato	GRO002	64	32	31	Sir Organic Farm	1984	0
8/17/2019	Potato	GRO003	57	18	18	Halder vegetarian	1026	0
3/5/2020	Tomato	GRO002	57	26	25	Kolkata daily bazar	1425	0
10/31/2019	Tomato	GRO002	79	23	21	Ladda vegetable centre	1659	0
8/28/2020	Onion	GRO001	65	50	49	Rupdasi market	3185	0
4/18/2021	Onion	GRO001	73	65	65	Tiwaree vee shoo	4745	0

This screenshot shows an Excel spreadsheet titled 'Data - Excel (Product Activation Failed)'. The spreadsheet contains a table with 9 columns: Creation_date, PO_item, Item_number, Quantity, Unit_max_price, Buying_price, Vendor_name, Total Buying Price, and Fraud. The data is organized into 29 rows, with the first row being the header. The 'Fraud' column contains values of 0 for all rows. The 'Total Buying Price' column contains values ranging from 1760 to 3185. The 'Buying_price' column contains values ranging from 14 to 65. The 'Quantity' column contains values ranging from 14 to 96. The 'Unit_max_price' column contains values ranging from 25 to 60. The 'Vendor_name' column contains various vendor names such as 'Garden fresh', 'puja vegetable shop', 'Freshbazar', 'Rupdasi market', 'Green Tokri', 'Abu variety stores', 'Newtown home delivery', 'Once more', 'Kolkata Mart', 'Kolkata daily bazar', 'RK wholesale vegetable', 'Newtown home delivery', 'Abu variety stores', 'Quadri vegetable shop', 'mohidul vegetable shop', 'Halder vegetarian', 'Kolkata daily bazar', 'Ladda vegetable centre', 'Rupdasi market', and 'Tiwaree vee shoo'. The 'PO_item' column contains values such as '1/19/2021 Onion', '2/14/2019 Tomato', '8/13/2019 Onion', '4/30/2021 Tomato', '5/13/2020 Onion', '6/21/2021 Onion', '11/4/2018 Tomato', '9/13/2020 Tomato', '2/3/2019 Onion', '7/14/2019 Tomato', '1/31/2021 Onion', '10/5/2020 Tomato', '9/29/2018 Onion', '10/16/2018 Onion', '2/23/2019 Onion', '1/10/2019 Tomato', '6/28/2021 Onion', '6/6/2021 Potato', '4/14/2019 Potato', '9/12/2020 Tomato', '1/26/2020 Onion', '3/8/2021 Onion', '8/8/2020 Tomato', '8/17/2019 Potato', '3/5/2020 Tomato', '10/31/2019 Tomato', '8/28/2020 Onion', and '4/18/2021 Onion'. The 'Item_number' column contains values ranging from 52 to 73. The 'Creation_date' column contains dates ranging from 1/19/2021 to 4/18/2021. The 'Quantity' column contains values ranging from 14 to 96. The 'Unit_max_price' column contains values ranging from 25 to 60. The 'Buying_price' column contains values ranging from 14 to 65. The 'Vendor_name' column contains various vendor names. The 'Total Buying Price' column contains values ranging from 1760 to 3185. The 'Fraud' column contains values of 0 for all rows. The spreadsheet is displayed in 'Sheet1' with a status bar at the bottom showing 'AVERAGE: 7658.912879 COUNT: 2124 SUM: 8087812'.

Creation_date	PO_item	Item_number	Quantity	Unit_max_price	Buying_price	Vendor_name	Total Buying Price	Fraud
1/19/2021	Onion	GRO001	52	60	58	Garden fresh	3016	0
2/14/2019	Tomato	GRO002	63	14	14	puja vegetable shop	882	0
8/13/2019	Onion	GRO001	63	30	30	Freshbazar	1890	0
4/30/2021	Tomato	GRO002	62	41	41	Rupdasi market	2542	0
5/13/2020	Onion	GRO001	53	45	44	Rupdasi market	2332	0
6/21/2021	Onion	GRO001	70	65	65	Green Tokri	4550	0
11/4/2018	Tomato	GRO002	51	11	10	Abu variety stores	510	0
9/13/2020	Tomato	GRO002	89	32	31	Newtown home delivery	2759	0
2/3/2019	Onion	GRO001	73	20	19	Rupdasi market	1387	0
7/14/2019	Tomato	GRO002	96	20	18	Once more	1728	0
1/31/2021	Onion	GRO001	92	60	56	Kolkata Mart	5152	0
10/5/2020	Tomato	GRO002	65	35	34	Kolkata daily bazar	2210	0
9/29/2018	Onion	GRO001	51	25	13	Rupdasi market	663	0
10/16/2018	Onion	GRO001	83	25	12	RK wholesale vegetable	996	0
2/23/2019	Onion	GRO001	88	20	20	Newtown home delivery	1760	0
1/10/2019	Tomato	GRO002	60	14	12	Abu variety stores	720	0
6/28/2021	Onion	GRO001	88	65	63	Quadri vegetable shop	5544	0
6/6/2021	Potato	GRO003	88	32	31	mohidul vegetable shop	2728	0
4/14/2019	Potato	GRO003	68	16	16	Abu variety stores	1088	0
9/12/2020	Tomato	GRO002	91	32	31	Tiwari veg shop	2821	0
1/26/2020	Onion	GRO001	96	40	37	Green Tokri	3552	0
3/8/2021	Onion	GRO001	65	60	60	Abu variety stores	3900	0
8/8/2020	Tomato	GRO002	64	32	31	Sir Organic Farm	1984	0
8/17/2019	Potato	GRO003	57	18	18	Halder vegetarian	1026	0
3/5/2020	Tomato	GRO002	57	26	25	Kolkata daily bazar	1425	0
10/31/2019	Tomato	GRO002	79	23	21	Ladda vegetable centre	1659	0
8/28/2020	Onion	GRO001	65	50	49	Rupdasi market	3185	0
4/18/2021	Onion	GRO001	73	65	65	Tiwaree vee shoo	4745	0

These are the raw attributes for given dataset(grocery dataset)analysis of different segments. Different types of business:

Dimensions:

- Product id-id of product purchased.
- Product item-Name of product purchased.
- Approving Officer-Name of officer who approved.
- Requesting Officer-Name of officer who send request for approval.
- Vendor Name-Name of vendor who purchased.

Measures:

- Total Buying price: It has total buying Price of product
- Buying Price-Buying price of specific product.
- Unit Max Price-Max price of product items.
- Creation_date-The date in which product is purchased.
- Fraud-Number of fraud happened in 0's and 1's

Exploratory Data Analysis(EDA):

Total Transactions
30,000

Total Expenditure
\$61.26M

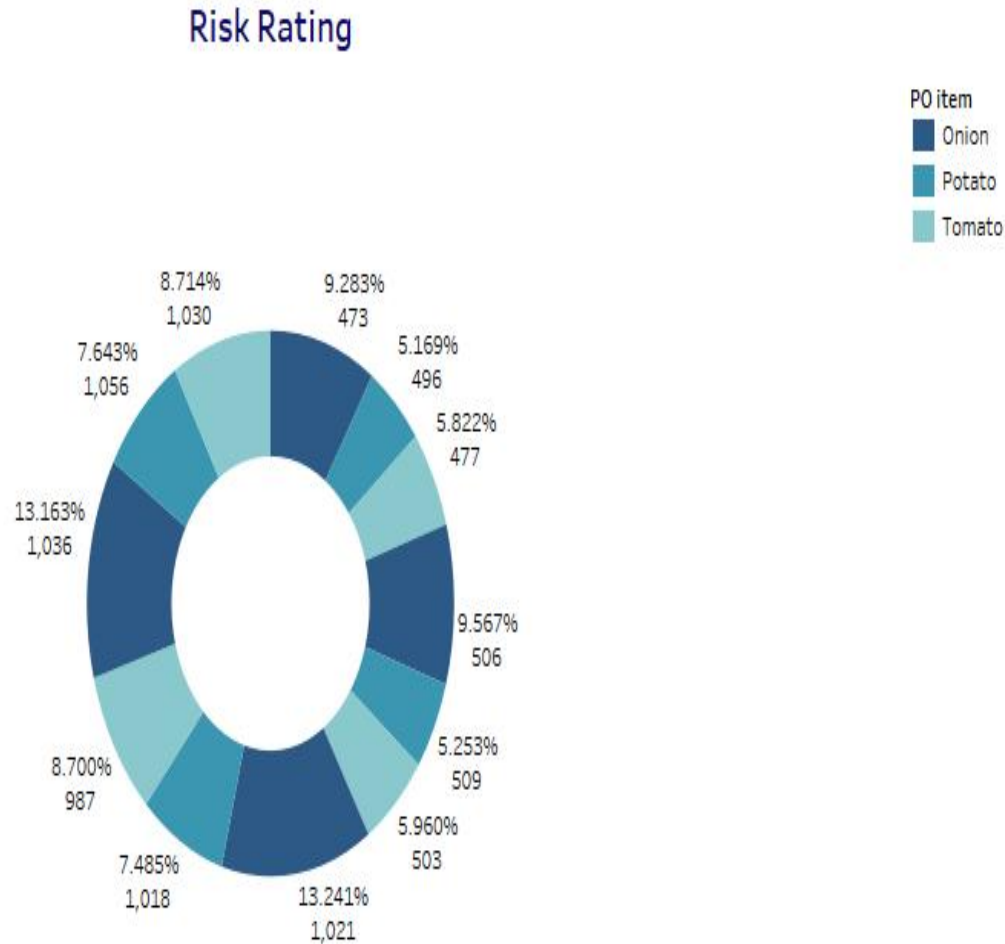
Quantity
\$2246.30K

AVERAGE AMOUNT
\$2042

Key inference:

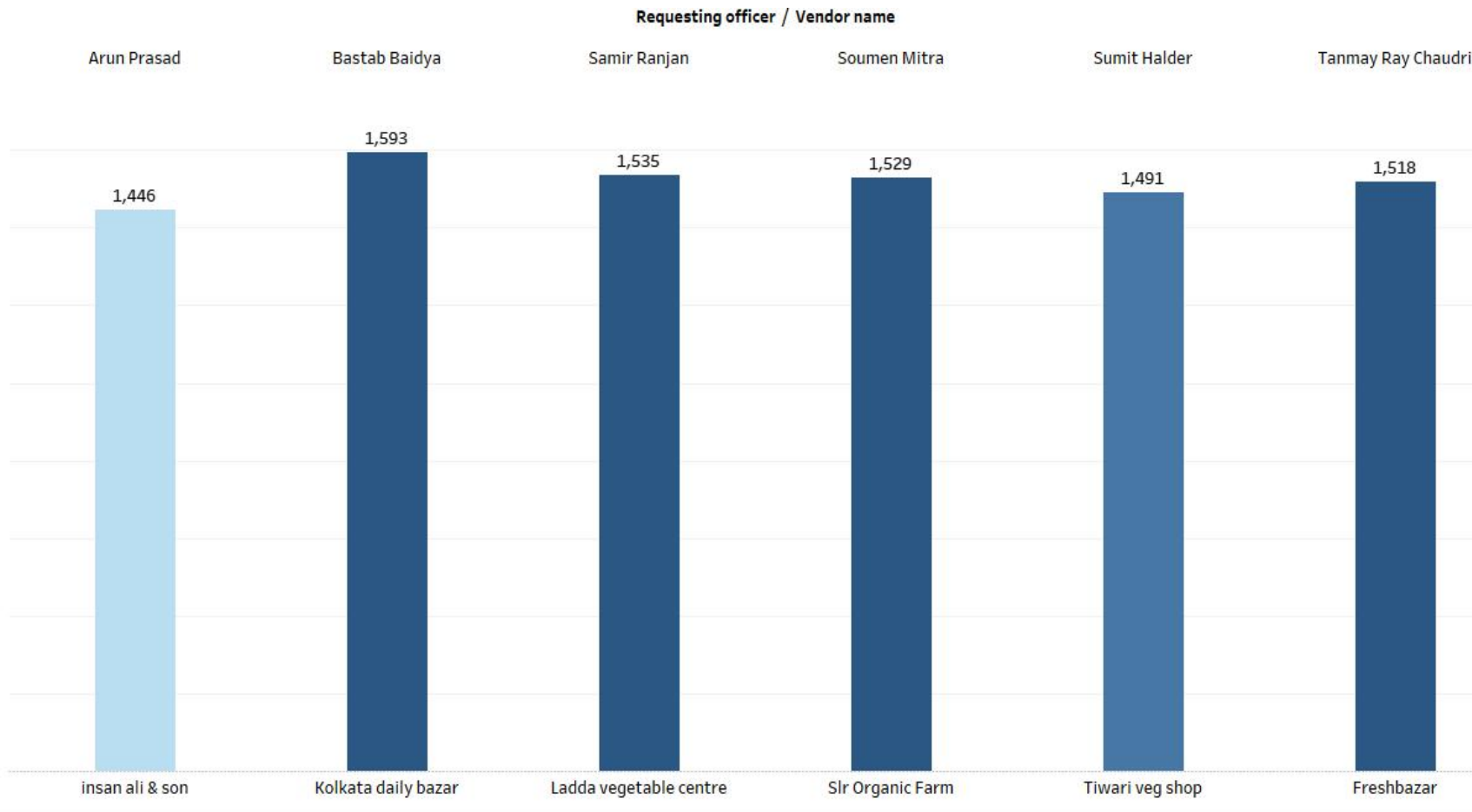
- From Marketing Data counted Sum of Total Buying Price, Sum of Quantity, Sum of Fraud and Total Transactions.
- The data is filtered on Creation date Year and PO item. The Creation date Year filter keeps 2018, 2019, 2020 and 2021. The PO item filter keeps Onion, Potato and Tomato.

EDA Cont..



- Key inference:**
- Color shows details about PO item. The marks are labeled by % of Total Buying Price and sum of Fraud. Details are shown for Approving officer and PO item. The data is filtered on Creation date, which ranges from 7/1/2018 to 6/29/2021. The view is filtered on Exclusions (Approving officer, PO item), which keeps 12 members.

Requesting Officers & Vendors Involved in Fraud:



Key inference:

- Sum of Fraud for each Vendor name broken down by Requesting officer. Color shows sum of Fraud. The view is filtered on Requesting officer, which keeps 6 of 20 members.

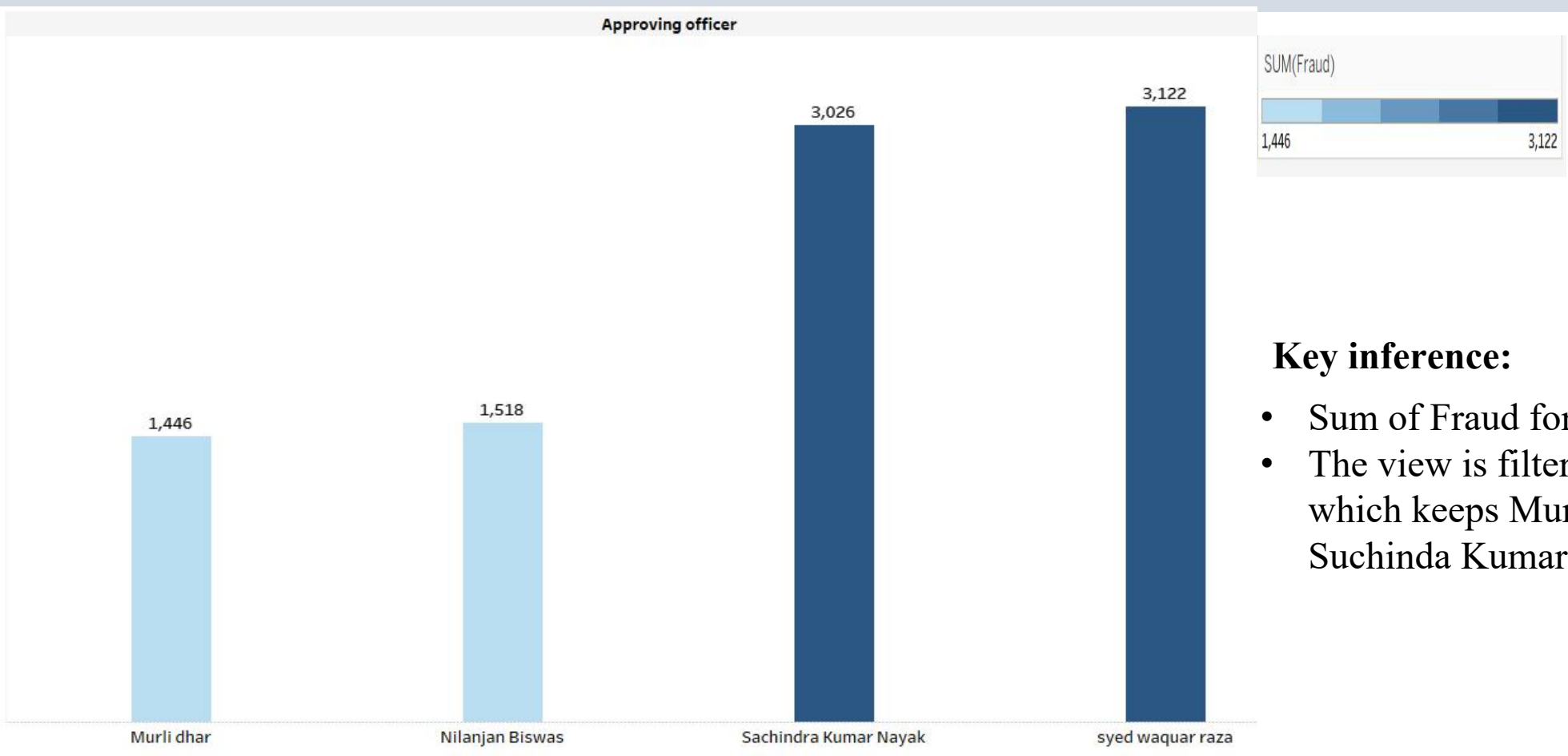
Year wise Calculations:

			Creation date			
Vendor name	PO item		2020	2019	2021	2018
Freshbazar	Onion	Fraud	174	161	86	85
		Count of Sheet1	263	247	128	126
		Buying price	11,184	8,761	5,967	3,891
		Unit max price	11,365	8,910	6,040	4,190
		Quantity	20,286	18,470	9,618	9,637
		Total Buying Price	862,673	652,173	450,107	297,109
	Potato	Fraud	165	161	96	87
		Count of Sheet1	250	244	135	149
		Buying price	5,476	4,798	3,235	2,404
		Unit max price	5,524	4,838	3,252	2,444
		Quantity	18,371	18,603	10,121	11,032
		Total Buying Price	401,176	363,800	239,865	177,891
	Tomato	Fraud	192	155	74	82
		Count of Sheet1	273	248	125	122
		Buying price	7,211	5,171	3,633	2,356
		Unit max price	7,278	5,263	3,688	2,407
		Quantity	20,201	18,681	9,365	9,004
		Total Buying Price	531,834	391,462	269,342	174,335
Kolkata daily bazar	Onion	Fraud	171	168	99	79
		Count of Sheet1	247	263	141	125
		Buying price	10,297	9,134	6,325	3,453
		Unit max price	10,430	9,335	6,400	3,820
		Quantity	18,609	20,125	10,865	9,590
		Total Buying Price	778,705	698,943	483,549	264,857
	Potato	Fraud	200	174	81	87
		Count of Sheet1	288	275	126	129
		Buying price	6,424	5,480	3,154	2,220
		Unit max price	6,474	5,530	3,204	2,270

Key inference:

- Buying price, Fraud, Quantity, count of Sheet1, Total Buying Price and Unit max price broken down by Creation date Year vs. Vendor name and PO item.
- Color shows Buying price. The view is filtered on Vendor name, which keeps Fresh bazar, Kolkata daily bazar, Ladda vegetable center, Slr Organic Farm and Tiwari veg shop.

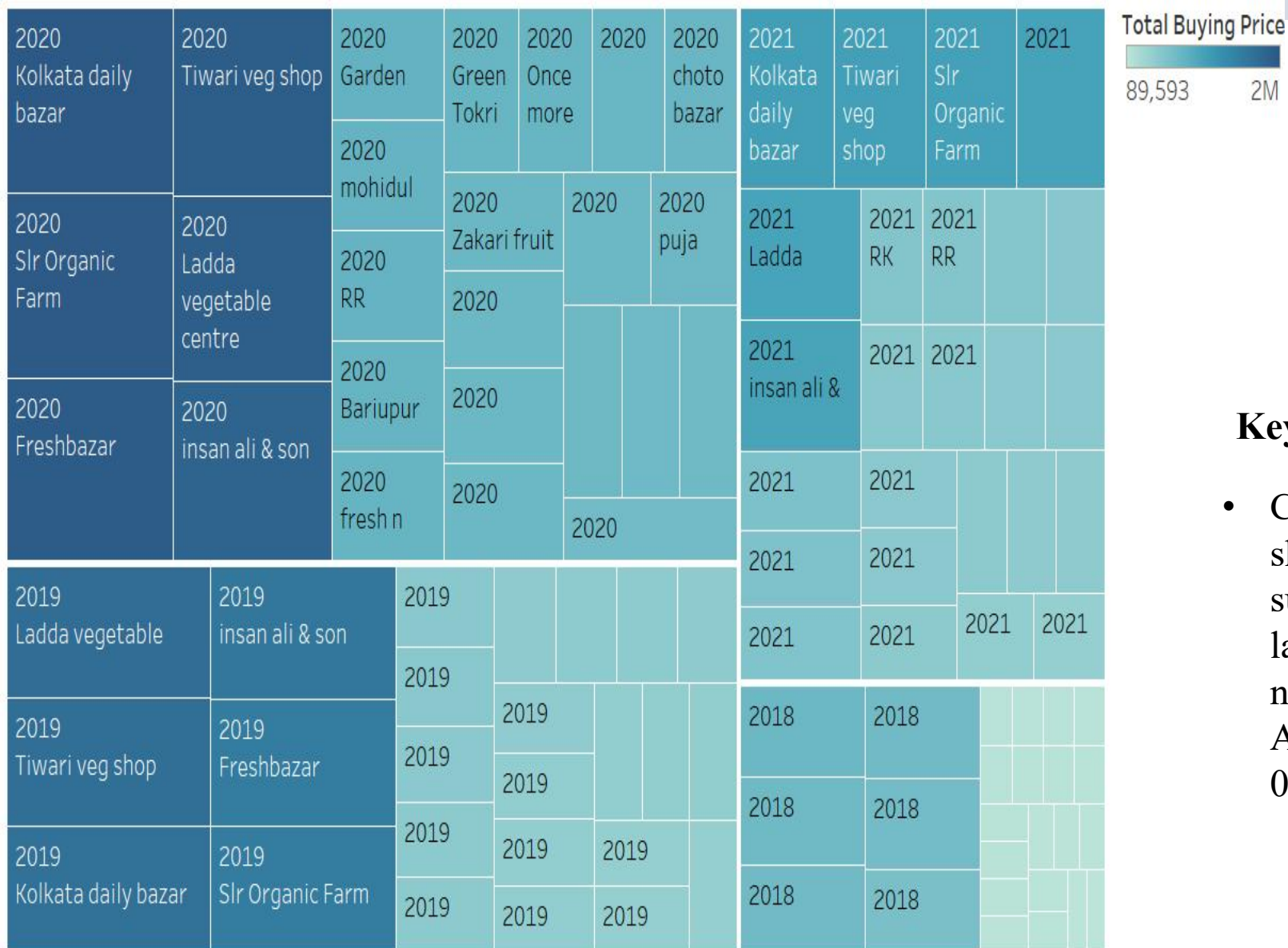
Approving Officers Involved in Fraud:



Key inference:

- Sum of Fraud for each Approving officer
- The view is filtered on Approving officer, which keeps Murli Dhar, Niranjan Biswas, Suchinda Kumar Nayak and Syed Waqar Raza.

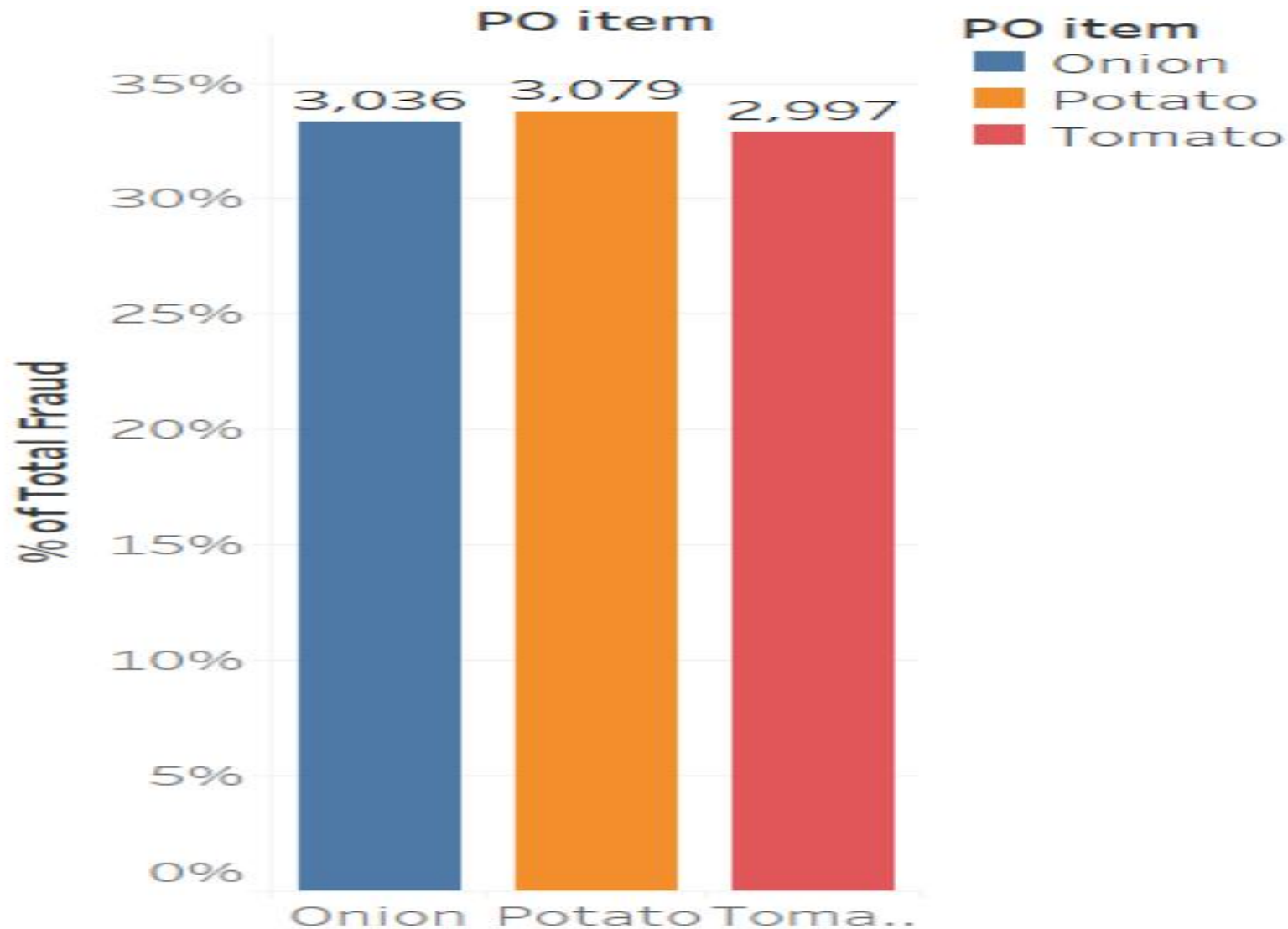
Total Buying Price Vs Vendor Names:



Key inference:

- Creation date year and vendor name. Color shows sum of Total Buying Price. Size shows sum of Total Buying Price. The marks are labeled by Creation date and Year and Vendor name. The details is filtered on Creation date Action. The Creation date ranges from 01/07/2018 to 29/06/20121.

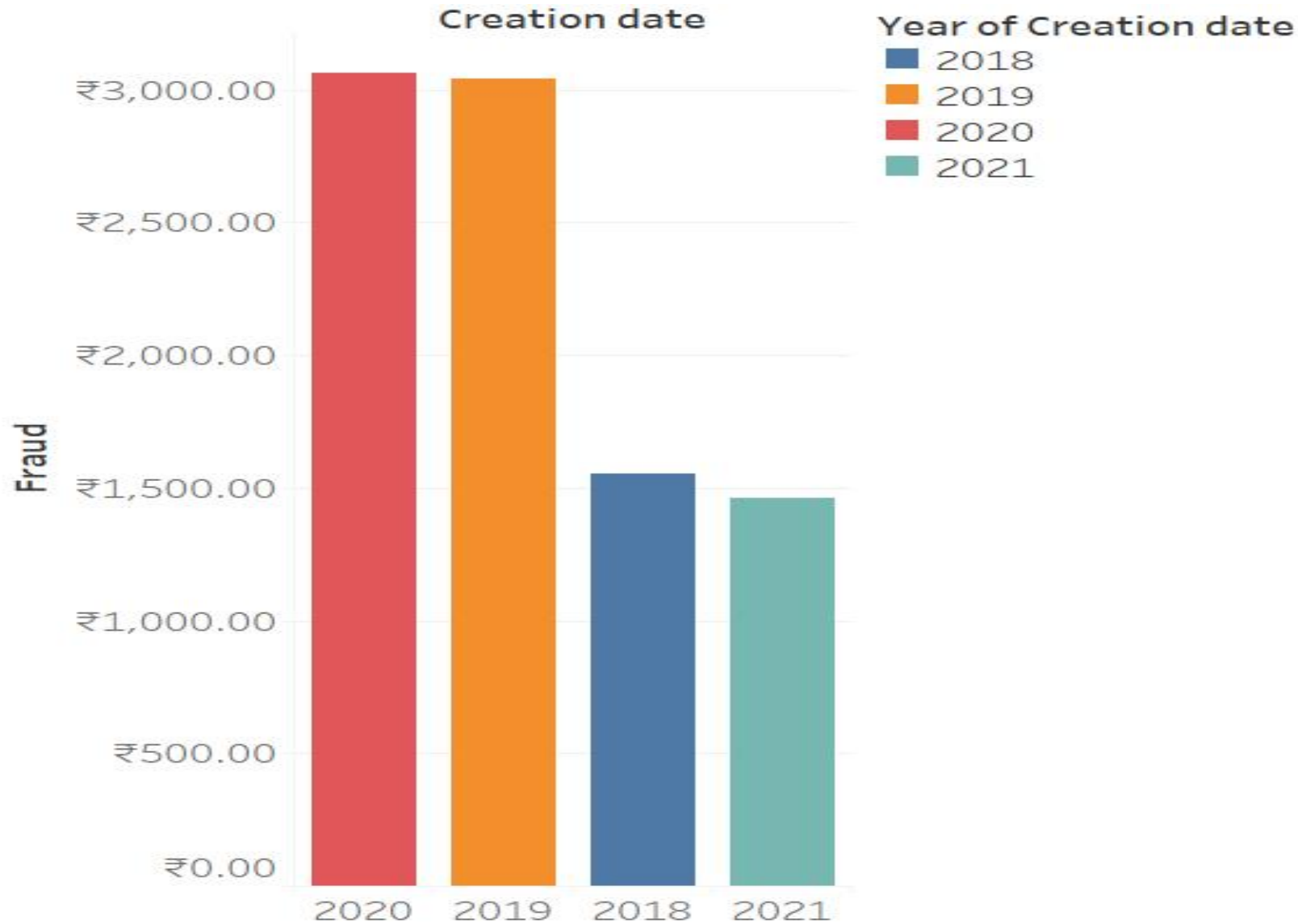
% Total Fraud vs Product Items:



Key inference:

- % of Total Fraud for each PO item. Color shows details about PO item. The marks are labeled by sum of Fraud.

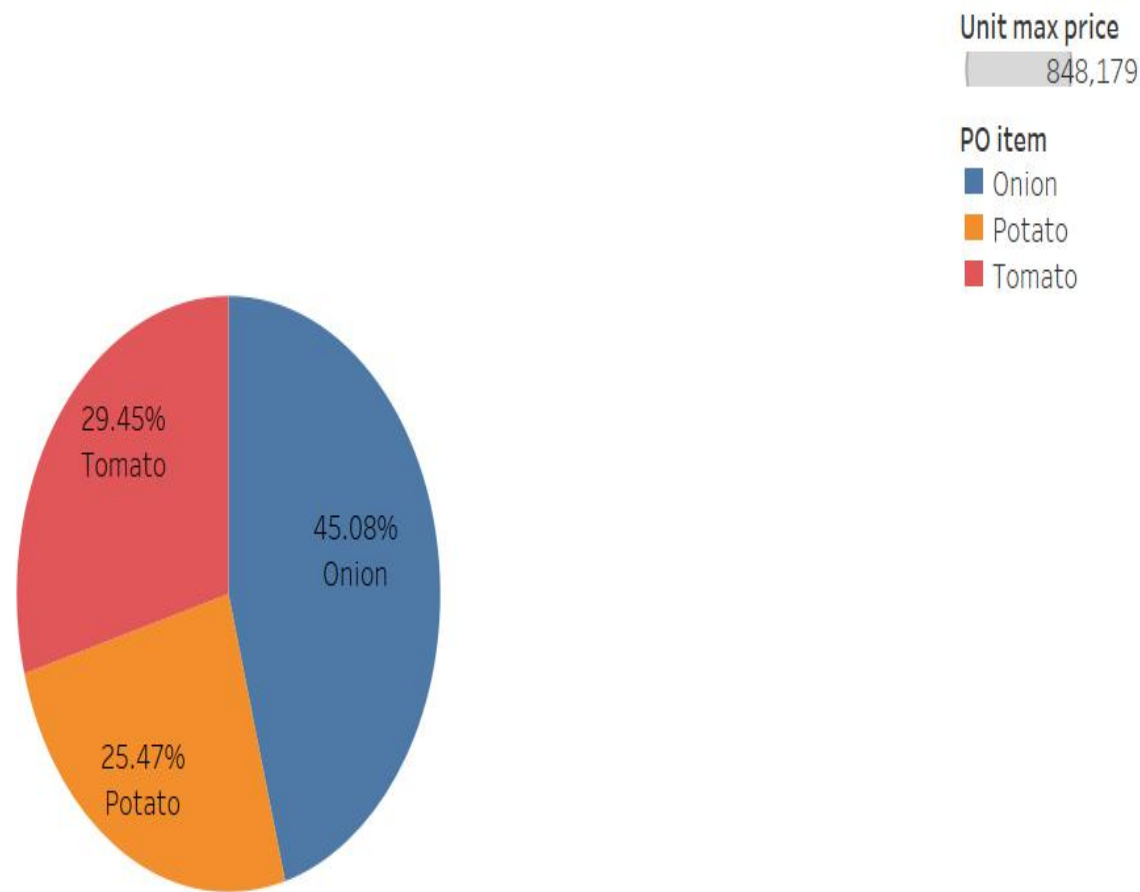
Fraud VS Year wise:



Key inference:

- Sum of Fraud for each Creation date Year. Color shows details about Creation date Year. The date is filtered on Creation date and Action. The Creation date filter ranges from 01/07/2018 to 29/06/2021

% Total Buying Price vs Product Items:



Key inference:

- % of Total Buying Price and PO item. Color shows details about po item. Size shows sum of Unit max price. The marks are labeled by % Total Price and PO item. The data is filtered on Creation date and Action. The Creation date filter keeps 1,095 of 1,095 members.

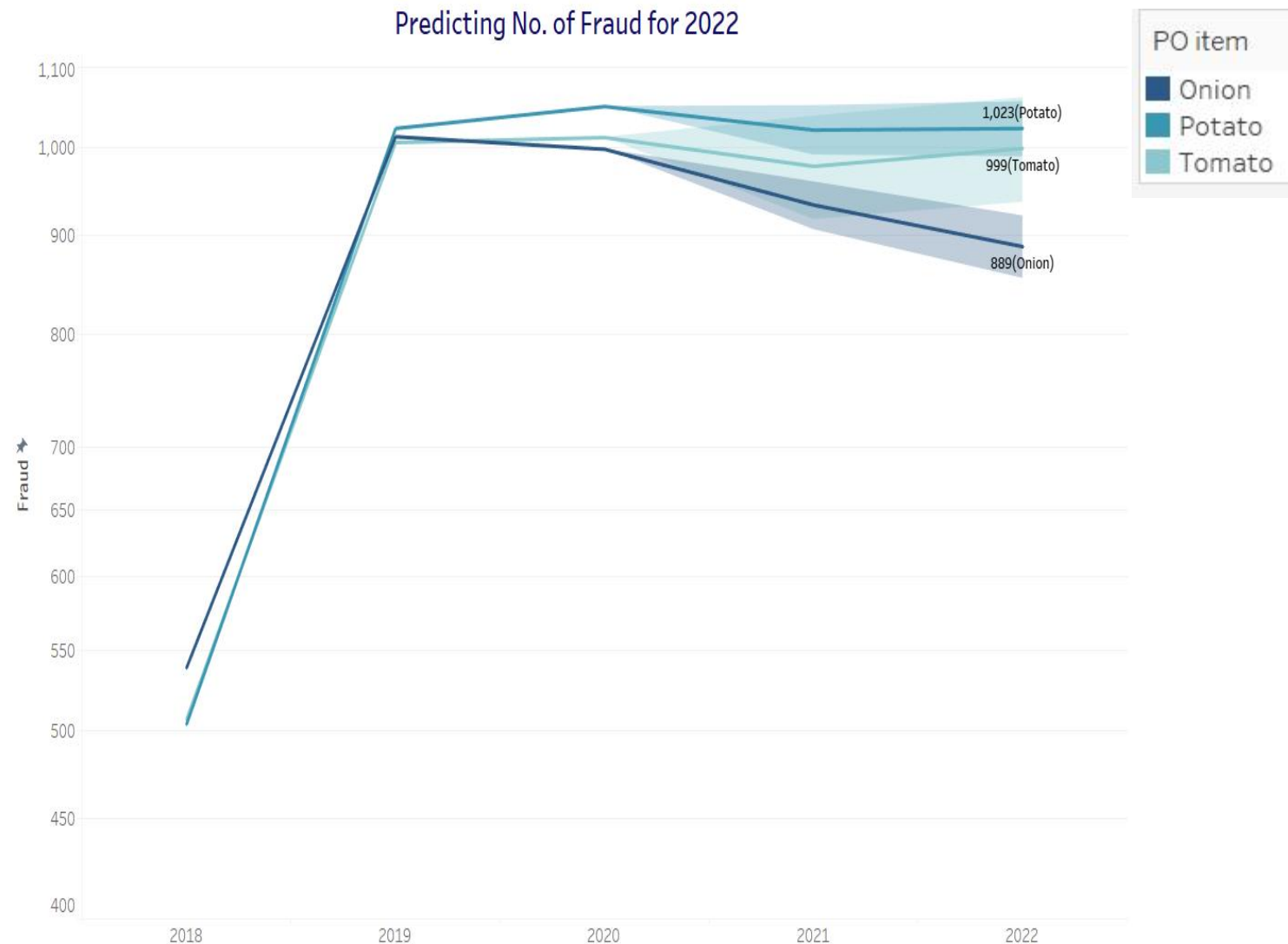
Officer's Fraud Report:

Requesting officer	Approving officer	Vendor name	Creation date			
			2018	2019	2020	2021
Bastab Baidya	syed waquar raza	Kolkata daily bazar	Rs. 549,602	Rs. 1,102,520	Rs. 1,153,863	Rs. 580,796
Samir Ranjan	Sachindra Kumar Nayak	Ladda vegetable centre	Rs. 586,854	Rs. 1,153,548	Rs. 1,026,952	Rs. 485,331
Tanmay Ray Chaudri	Nilanjan Biswas	Freshbazar	Rs. 545,572	Rs. 1,009,943	Rs. 1,141,841	Rs. 528,410
Soumen Mitra	syed waquar raza	Slr Organic Farm	Rs. 559,192	Rs. 1,016,832	Rs. 1,129,397	Rs. 485,662
Sumit Halder	Sachindra Kumar Nayak	Tiwari veg shop	Rs. 485,001	Rs. 1,120,716	Rs. 1,054,827	Rs. 497,678
Arun Prasad	Murli dhar	insan ali & son	Rs. 525,358	Rs. 1,070,377	Rs. 991,045	Rs. 446,796
Grand Total			Rs. 3,251,579	Rs. 6,473,936	Rs. 6,497,925	Rs. 3,024,673

Key inference:

- Sum of Total Buying Price broken down by Creation date Year vs. Requesting officer, Approving officer and Vendor name. Color shows details about Category. The data is filtered on Year and PO item.

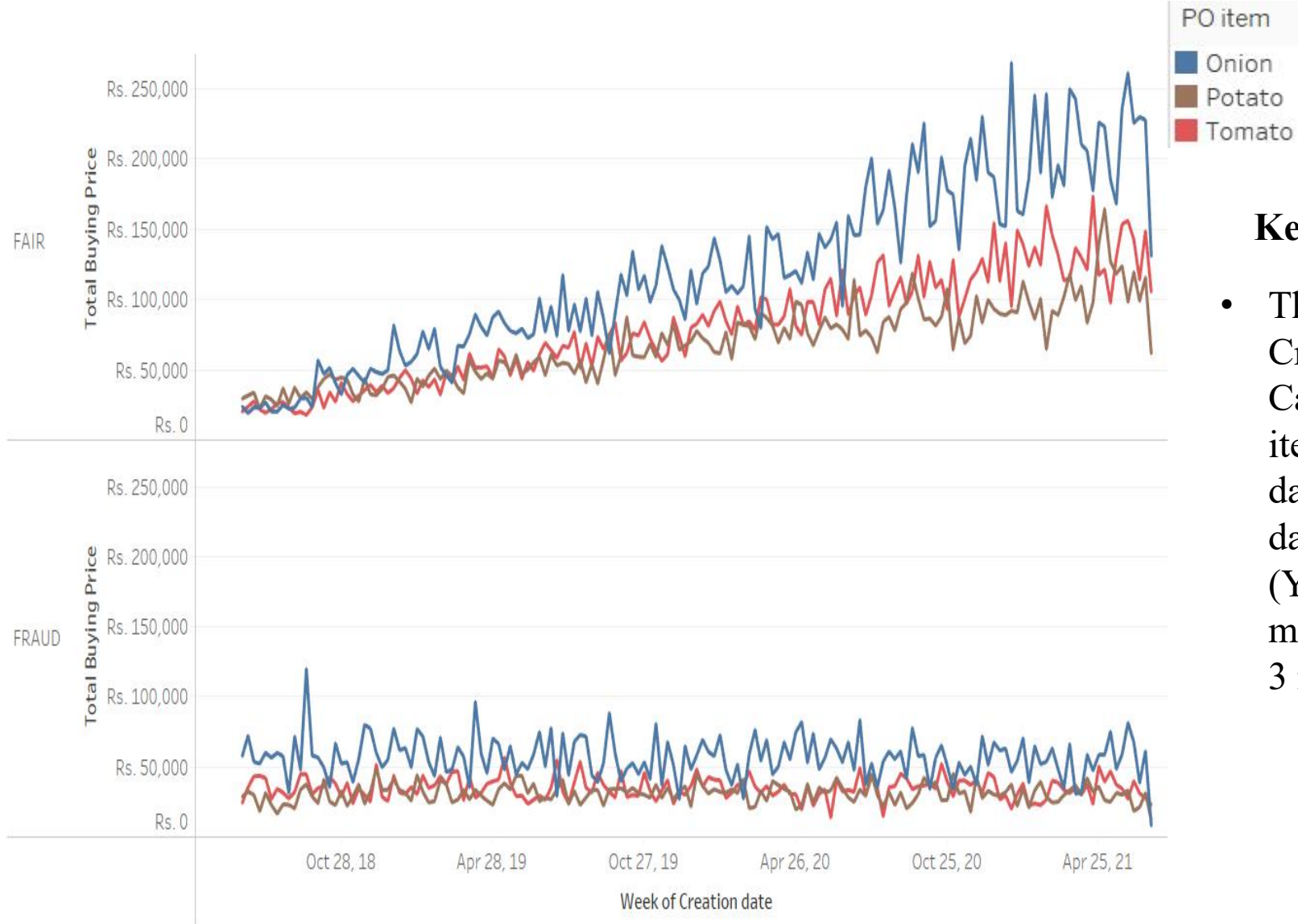
Forecasting:



Key inference:

- The trend of sum of Fraud (actual & forecast) for Creation date Year. Color shows details about PO item. The marks are labeled by sum of Fraud (actual & forecast) and PO item.

Trend Lines:



Key inference:

- The trend of sum of Total Buying Price for Creation date Week broken down by Category. Color shows details about PO item. Details are shown for PO item. The data is filtered on Action (YEAR(Creation date)) and Action (PO item). The Action (YEAR(Creation date)) filter keeps 4 members. The Action (PO item) filter keeps 3 members.

Dashboard:



Procurement Fraud Transactions

Item Name

All

YEAR

All

30,000
Total Transactions

2,246.30K
Quantity

\$61.26M
Total Expenditure

\$2,042
Average Amount

Onion Potato Tomato

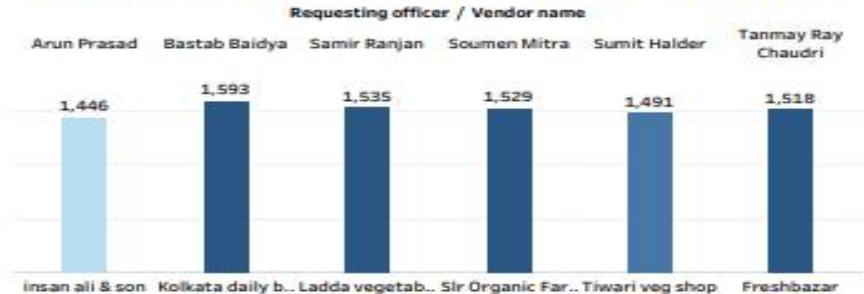
1,446 1,518

1,446 3,122

Risk Rating



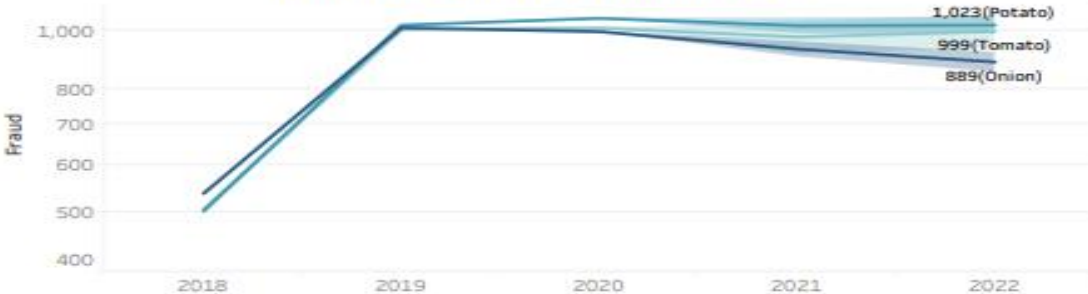
Requesting Officers & Vendors Involved in Fraud



Approving Officers Involved in Fraud

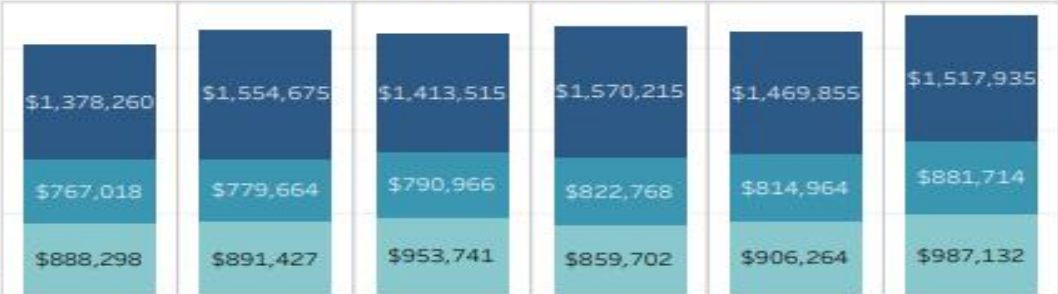


Predicting No. of Fraud for 2022



Fraud Transaction Details

PS: More Details on ToolTip...



Inferences(Conclusion And Future Scope):

Conclusion:

- In this project, we analyzed that total number of fraud transaction done from year 2018 to 2021 is 9112 and total buying price is 61258k.
- The visualization is done and ranking is given to most number of fraud transaction done officer and also the fraud value with date.
- The total number of buying price of onion is high compared to potato and tomato.
- Approving officer Syed Waqar Raza, Requesting officer Bastab Babayan and Kolkata daily bazar involved Highest No.of (542,534) Fraud Transactions from Total Buying cost \$881.74k, \$987.13k w.r.t. selected items Potato, Tomato for Total years.
- We Detected Risk Assessment and Total Percentage of Fraud Involved to Total Years.

Future Scope:

- Forecasting technique predict the future requirements of product items.
- Forecasting method Estimates 1,023 for Potato, 999 for Tomato, 889 for Onion Fraud Transactions estimated for 2022 Year.
- We also forecast the fraud in respective year, product item and total buying price is increasing with respect to fraud.
- Forecasting techniques allows companies to minimize fraud in the respective years.



THANK you