Adverse Side Effects of Vaccines, 2006-2016

Adriana Cardenas & Sofia Shubert

The Existing Problem

- There is misinformation/fear-mongering about the safety of vaccines in the United States
- Three in ten Americans believe vaccines should not be mandatory; 9% believe the measles vaccine is not safe
- Diseases that were once eradicated such measles, mumps, and whooping cough

Sources:
http://time.com/27308/4-diseases-making-a-comeback-thanks-to-anti-vaxxers/
https://measlesrubellainitiative.org/anti-vaccination-movement/

Our Questions

- Which vaccines cause adverse side effects considered severe/life-threatening and at what amounts?

- Was there a spike in severe side effects such as death or disablement in any years between 2006-2016?

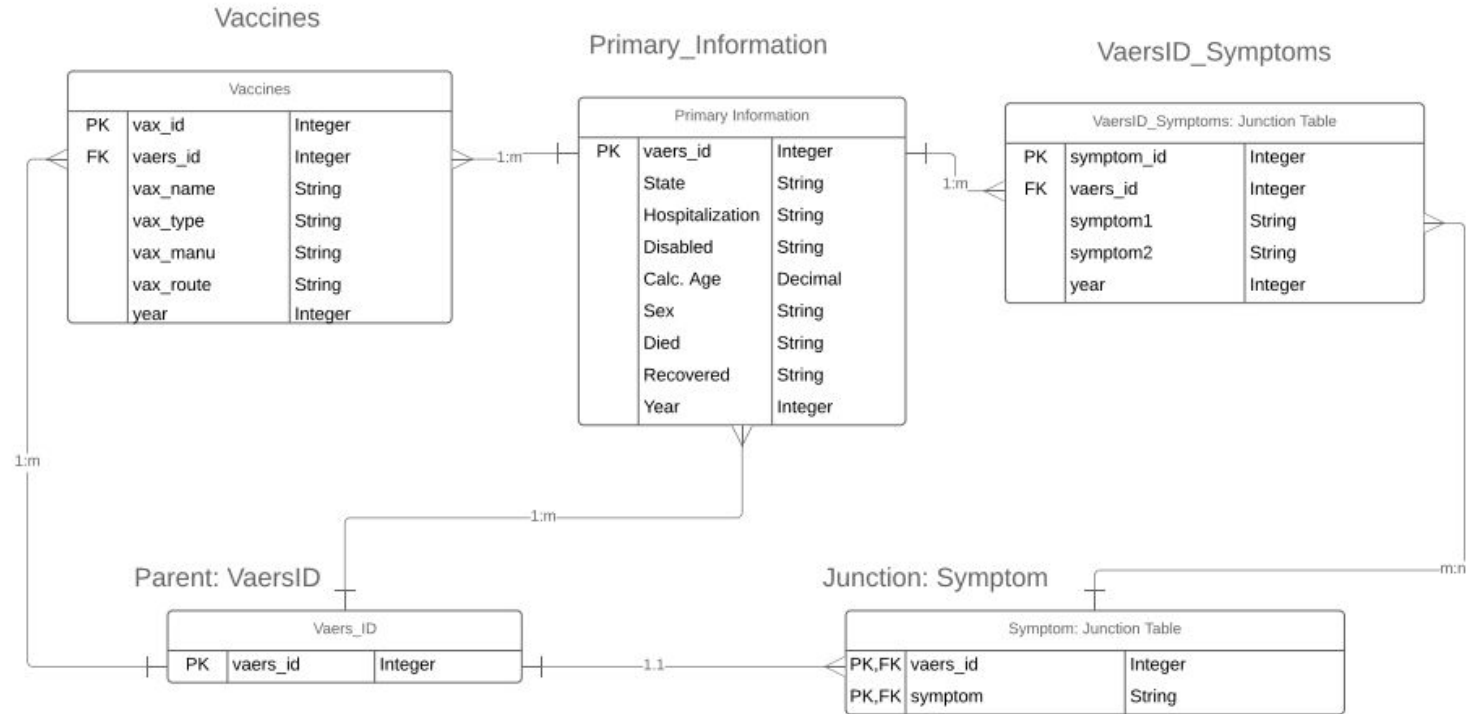- Which manufacturers had the most reports made for their products?

Dataset 1 & 2

From **CDC - Centers for Disease Control - Vaccine Adverse Event Reporting System** for the years 2006-2016:

1. Vaccine Adverse Event Reporting System - Symptoms
   Vaccine Adverse Event Reporting System - Vaccines

2. Vaccine Adverse Event Reporting System - Primary Information

Final ERD

Vaccines

	Vaccines	
PK	vax_id	Integer
FK	vaers_id	Integer
	vax_name	String
	vax_type	String
	vax_manu	String
	vax_route	String
	year	Integer

Primary_Information

	Primary Information	
PK	vaers_id	Integer
	State	String
	Hospitalization	String
	Disabled	String
	Calc. Age	Decimal
	Sex	String
	Died	String
	Recovered	String
	Year	Integer

VaersID_Symptoms

	VaersID_Symptoms: Junction Table	
PK	symptom_id	Integer
FK	vaers_id	Integer
	symptom1	String
	symptom2	String
	year	Integer

1:m

1:m

1:m

1:m

m:n

1.1

Parent: VaersID

	Vaers_ID	
PK	vaers_id	Integer

Junction: Symptom

	Symptom: Junction Table	
PK,FK	vaers_id	Integer
PK,FK	symptom	String

Our Solution

Live demo including:

- BigQuery Views
- Cross-dataset SQL Queries
- Data Visualizations

Transforms: DAG/Airflow & Beam (github)

```python
with models.DAG(
        'workflow',
        schedule_interval=datetime.timedelta(days=1),
        default_args=default_dag_args) as dag:

    fillPrimaryInfo = BashOperator(
        task_id='fill_workflow',
        bash_command='bq query --use_legacy_sql=false "' + sql1 + '"')  # default trigger rule is all_success

    changeNullsState = BashOperator(
        task_id='change_nulls_to_false_state',
        bash_command='bq query --use_legacy_sql=false "' + sql2 + '"')  # default trigger rule is all_success

    changeNullsHospital = BashOperator(
        task_id='change_nulls_to_false_hospitalization',
        bash_command='bq query --use_legacy_sql=false "' + sql3 + '"')  # default trigger rule is all_success

    changeNullsDisabled = BashOperator(
        task_id='change_nulls_to_false_disabled',
        bash_command='bq query --use_legacy_sql=false "' + sql4 + '"')  # default trigger rule is all_success

    changeNullsDied = BashOperator(
        task_id='change_nulls_to_false',
        bash_command='bq query --use_legacy_sql=false "' + sql5 + '"')  # default trigger rule is all_success

    changeNullsRecovered = BashOperator(
        task_id='change_nulls_to_u',
        bash_command='bq query --use_legacy_sql=false "' + sql6 + '"')  # default trigger rule is all_success

    normalizeState = BashOperator(
        task_id='normalize_uncapitalized_values_state',
        bash_command='bq query --use_legacy_sql=false "' + sql7 + '"')  # default trigger rule is all_success

    fillPrimaryInfo >> changeNullsState >> changeNullsHospital, changeNullsDisabled, changeNullsDied, changeNullsRecovered, normalizeState
```

https://github.com/cs327e-spring2019/ssac/blob/master/transform_personalinfo_single.py

Conclusions/Future Improvements

- Merck & Co. Inc. had the most deaths associated with it
- Infants (43%) and the elderly (17.8%) were age groups most impacted by deaths
- Reports of adverse side effects spiked in 2007, 2009 and 2015
- Original data was subjective due to patients likely self reporting their symptoms or issues with correlation versus causation. Final conclusions are not necessarily reflective of the safety of vaccines
- In a future study, state regulations on whether vaccines are mandatory or not should be taken into consideration.
- Company size and production should be taken into account when calculating proportions of adverse reactions