

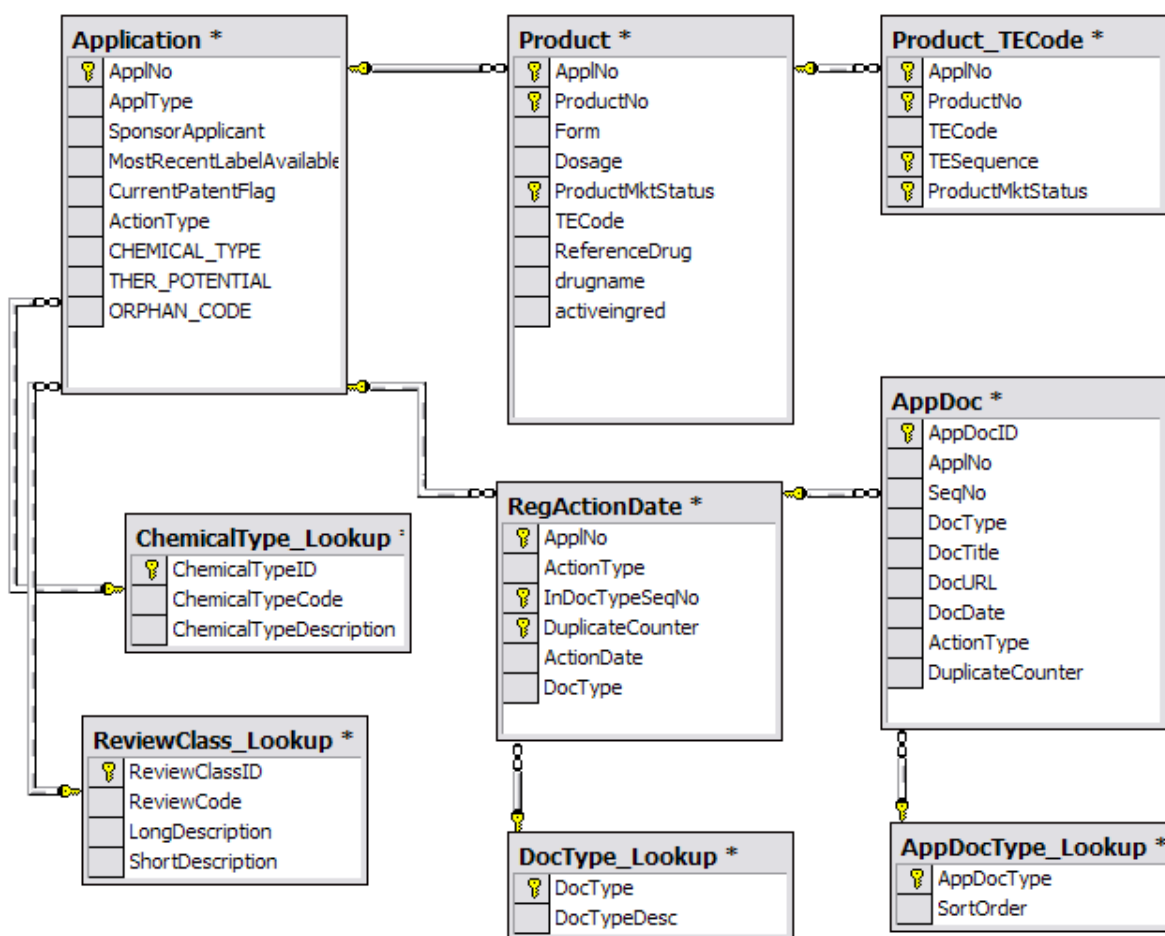
## Project: SQL Data Analysis and Visualization with Power BI for FDA

The **U.S. Food and Drug Administration (FDA)** is a federal agency responsible for safeguarding public health in the United States. It oversees various areas, including food safety, pharmaceuticals, medical devices, cosmetics, tobacco, and veterinary products. The **FDA** evaluates the safety and efficacy of drugs, biological products, and medical devices, conducts inspections, and enforces regulations. It also regulates food production and distribution, conducts research, provides public health education, and monitors and responds to emerging health risks. The **FDA's** mission is to protect and promote public health by ensuring the safety and effectiveness of a wide range of products and substances. **iVision** is collaborating with **FDA** to perform Data analysis on **FDA's** dataset. **FDA** has provided **iVision** with **SQL script** of their dataset. As an analyst at **iVision** your role is to create informative reports by conducting a thorough analysis of the data using powerful tools like MySQL and Power BI.

### Software/Tools:

- MySQL
- Power BI

### Data Schema (Draugs@FDA Dataset):



---

## Project Objectives

### Part - 1 SQL Queries:

#### **Task 1: Identifying Approval Trends**

1. Determine the number of drugs approved each year and provide insights into the yearly trends.
2. Identify the top three years that got the highest and lowest approvals, in descending and ascending order, respectively.
3. Explore approval trends over the years based on sponsors.
4. Rank sponsors based on the total number of approvals they received each year between 1939 and 1960.

#### **Task 2: Segmentation Analysis Based on Drug MarketingStatus**

1. Group products based on MarketingStatus. Provide meaningful insights into the segmentation patterns.
2. Calculate the total number of applications for each MarketingStatus year-wise after the year 2010.
3. Identify the top MarketingStatus with the maximum number of applications and analyze its trend over time.

#### **Task 3: Analyzing Products**

1. Categorize Products by dosage form and analyze their distribution.
2. Calculate the total number of approvals for each dosage form and identify the most successful forms.
3. Investigate yearly trends related to successful forms.

#### **Task 4: Exploring Therapeutic Classes and Approval Trends**

1. Analyze drug approvals based on therapeutic evaluation code (TE\_Code).
2. Determine the therapeutic evaluation code (TE\_Code) with the highest number of Approvals in each year.

### Part - 2 Power BI Visualizations:

- i. Visualize the yearly approval trends of drugs. Highlight any significant patterns and/or fluctuations, if any.
- ii. Explore approval trends over the years based on different sponsors. Uncover patterns and changes in approval rates among sponsors.
- iii. Visualize the segmentation of products based on MarketingStatus.
- iv. Show the total number of applications for each MarketingStatus. Enable users to filter by years and MarketingStatus for detailed analysis.

- 
- v. Analyze the grouping of drugs by dosage form. Visualize the distribution of approvals across different forms. Identify the most successful dosage form.
  - vi. Visualize drug approvals based on therapeutic classes. Identify classes with the highest number of approvals.

### **Artifacts to be generated (For Learners):**

- Power BI File (.pbix)
- SQL Script (.sql)
- Presentation (Slides) summarizing the project.
  - Learner needs to present this to Faculty (Viva-Voice)
  - Presentation time: 10 Mins
- Artifacts generated need to be submitted in vLearn on or before deadline.
- Create a zip file of Power BI File, SQL Script File and Presentation file.
  - Zip file name: <LearnerName>\_ABADS\_<Batch>.zip
  - E.g.: KartikMudaliar\_ABADS\_B10.zip

### **Grading Criteria (For Faculty & Learners):**

Learners will be graded on below mentioned parameters (20 points for each):

- 1) Relevance and approach of Analysis (SQL Query Output)
- 2) Quality of Output
- 3) Overall presentation
- 4) Response to Questions
- 5) Creativity & Innovation

**Total Points: 100**