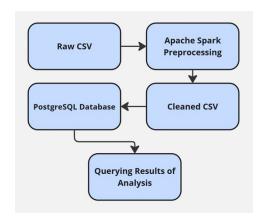
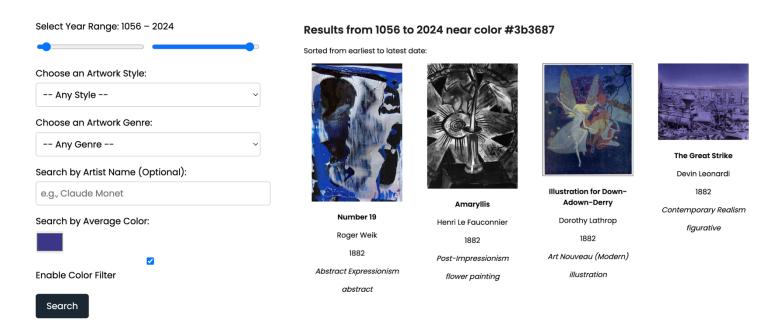
Built a full-stack data analysis pipeline to process and explore large datasets (Glassdoor job reviews, IMDB ratings, and Artworks metadata). Data was cleaned and processed with Spark, stored in MySQL, and exposed via a Flask API. Designed a responsive web interface (HTML, CSS, JS) to visualize trends and insights from millions of records.



Outline of the project pipeline

#### **ARTWORKS**



# **Glassdoor Analytics Dashboard**

Let's get some insights into company reviews and job ratings!!!



### **Job Analysis**

Ratings Filtered by Job:	
e.g., Software Engineer	
Top N Items for Job or Company Reviews:	
5	
Set Threshold for Average Rating:	
0.0	
Pros/Cons Ratio Order:	
Highest	~
Select Analyses:	
☐ Top N Jobs by Reviews	
☐ Top N most reviewed companies	
☐ Jobs by pros/cons word count ratio	
Jobs with average rating above threshold	
Show Results	

## Filtered Job: "manager"

Category: manager Review Count: 890581 Average Rating: 3.63

### **Top N Jobs by Reviews**

Job	Review Count
anonymous employee	1387461
None	563058
sales associate	127645
software engineer	108260

### **IMDB**



# **Movie Analysis**

