



14 DAYS

AI CHALLENGE

DAY 14

Topic:

AI-Powered Analytics: Genie & Mosaic AI

Challenge:

1. Use Genie to query data with natural language
2. Explore Mosaic AI features
3. Build simple NLP task
4. Create AI-powered insights

#DatabricksWithIDC

The **average revenue** generated by movies in each industry is **\$7,868.01** for Bollywood and **\$928.90** for Hollywood. Bollywood movies have a significantly higher average revenue compared to Hollywood movies in this dataset.

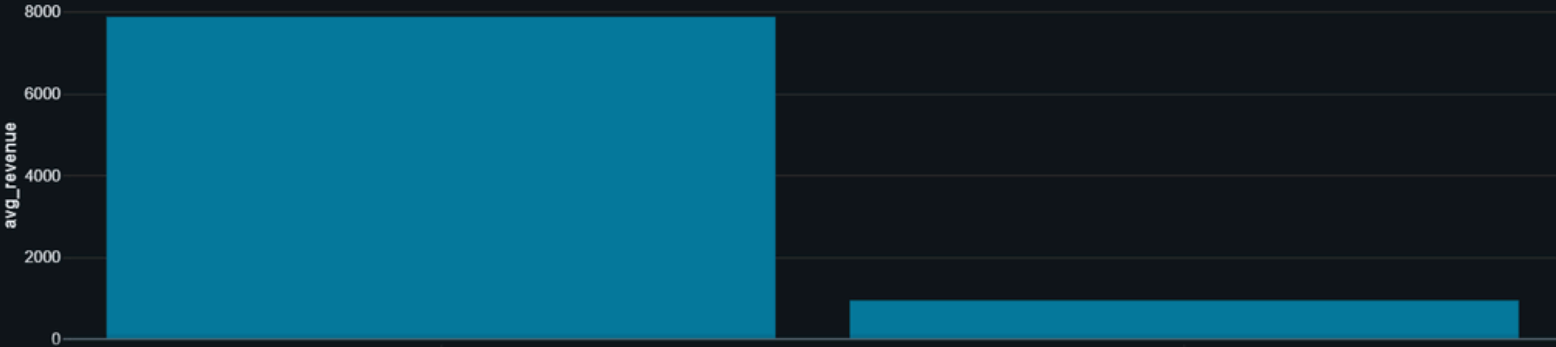
Result table (2 rows)

Download icon Hide code More options icon

```
1 SELECT
2   `industry`,
3   AVG(`revenue`) AS avg_revenue
4 FROM
5   `workspace`.`default`.`movies`
6 WHERE
7   `industry` IS NOT NULL
8   AND `revenue` IS NOT NULL
9 GROUP BY
10  `industry`;
```

1	Bollywood	7,868.005882
2	Hollywood	928.895

Average Revenue by Industry



+ Code

+ Text

✦ Assistant

▶ Run suggested

5

Python



summarize trends in movie ratings



Reject Esc

Accept ⌘ / Tab



```
# Assuming a Spark DataFrame 'movie_ratings' with columns: 'movie', 'rating', 'date'
from pyspark.sql import functions as F

# Aggregate average rating per year
ratings_trend = (
    movie_ratings
    .withColumn("year", F.year("date"))
    .groupBy("year")
    .agg(F.avg("rating").alias("avg_rating"), F.count("rating").alias("num_ratings"))
    .orderBy("year")
)

display(ratings_trend)
```

```
classifier = pipeline("sentiment-analysis")
reviews = [
    "Impressive", "Needs Improvement"
]

results = classifier(reviews)
with mlflow.start_run(run_name="Acting-analysis"):
    mlflow.log_param("model", "Acting-analysis")
    mlflow.log_metric("accuracy", 0.9)
    for i, result in enumerate(results):
        if "score" in result:
            mlflow.log_metric(f"score_{i}", result["score"])
```

▼ (1) MLflow run

Logged 1 run [↗](#) to an experiment [↗](#) in MLflow. [Learn more ↗](#)

Metrics (3)

Metric	Latest	Min	Max
accuracy	0.9	0.9	0.9
score_0	0.999863862991333	0.999863862991333	0.999863862991333
score_1	0.9966731071472168	0.9966731071472168	0.9966731071472168

Parameters (1)

Parameter	Value
model	Acting-analysis