

Data Sources

Real-time streams:

- **Stream 1:** Ad Click Stream (*Refer the sample data and write script to generate mock data with real time frequency*)

- Fields: ad_id, user_id, click_timestamp, platform, location

```
{
  "ad_id": 1001,
  "user_id": "U1234",
  "click_timestamp": "2023-09-09 10:00:00",
  "platform": "Web",
  "location": "New York"
}
{
  "ad_id": 1002,
  "user_id": "U5678",
  "click_timestamp": "2023-09-09 10:05:23",
  "platform": "Mobile",
  "location": "Los Angeles"
}
{
  "ad_id": 1003,
  "user_id": "U9101",
  "click_timestamp": "2023-09-09 10:10:34",
  "platform": "Tablet",
  "location": "San Francisco"
}
```

- **Stream 2:** Ad Conversions Stream (*Refer the sample data and write script to generate mock data with real time frequency*)

- Fields: ad_id, user_id, conversion_timestamp, product_id, revenue

```
{
  "ad_id": 1001,
  "user_id": "U1234",
  "conversion_timestamp": "2023-09-09 10:20:00",
  "product_id": "P501",
  "revenue": 50.00
}
{
  "ad_id": 1002,
  "user_id": "U5678",
  "conversion_timestamp": "2023-09-09 10:30:23",
  "product_id": "P502",
  "revenue": 75.00
}
```

- **Campaigns** (*Slowly changing dimension table which should be loaded beforehand in downstream system, refer sample data*):

- Fields: ad_id, campaign, product, target_start_date, target_end_date

ad_id, campaign, product, target_start_date, target_end_date

1001, Back-to-School, Laptop, 2023-08-01, 2023-10-31

1002, Summer Sale, Smartphone, 2023-07-15, 2023-09-15

1003, Fall Fashion, Jacket, 2023-09-01, 2023-11-30

● **User Demographics** (*Slowly changing dimension table which should be loaded beforehand in downstream system, refer sample data*):

- Fields: user_id, age, gender, interests

user_id, age, gender, interests

U1234, 22, Male, Music;Travel

U5678, 30, Female, Reading;Photography

U9101, 27, Male, Gaming;Cooking