

CHRIST (Deemed to be University), Bangalore - 560029

Department of Statistics and Data Science

Neekitasingh Rajput

2441339

Class: 3DM

Course: DSC261-3 - Data Visualization

Date: 28/07/2025

Exercise No: Lab Assignment-6

Data Joining & Data Blending in Tableau

Data Joining

Data joining is the process of merging two or more tables based on a common field (called a key). This technique is used when the tables are from the same data source and have a direct relationship. In Tableau, we can perform different types of joins like Inner Join, Left Join, Right Join, and Outer Join depending on how we want to combine the data.

In this task, data joining helps us to create a unified view of related information, such as connecting patients to their appointments or doctors to their schedules.

Data Blending

Data blending is used when data resides in different sources or when tables are not directly joinable in Tableau. It creates a relationship between primary and secondary data sources by linking common dimensions. Unlike joining, blending happens after data aggregation, which makes it suitable for loosely related data.

In this task, blending allows us to relate appointment-level data with billing-level data to derive insights across different scopes of information.

Healthcare Dataset

About Dataset-

The dataset used for this task simulates a healthcare management system and is organized across five interconnected sheets: appointments, patients, doctors, billing, and treatments. It captures key aspects of hospital operations, including patient demographics, doctor details, appointment scheduling, billing information, and treatment records. Each sheet contains structured data that, when combined, allows for meaningful analysis of clinical activities and financial

performance. The dataset provides a realistic foundation for practicing data joining and blending techniques in Tableau, enabling the extraction of insights across different domains of healthcare service delivery.

appointments

Attribute	Description
appointment_id	Unique ID for each appointment (Categorical)
patient_id	ID referencing the patient (Categorical)
doctor_id	ID referencing the doctor (Categorical)
appointment_date	Date of the appointment (Date)
department	Medical department (e.g., Cardiology) (Categorical)
status	Status of the appointment (Completed, Cancelled, etc.) (Categorical)

patients

Attribute	Description
patient_id	Unique ID for each patient (Categorical)
name	Full name of the patient (Text/String)
age	Age in years (Numerical)
gender	Gender (Male, Female, etc.) (Categorical)
contact_number	Patient's phone number (Text/String)

doctors

Attribute	Description
doctor_id	Unique ID for each doctor (Categorical)
name	Doctor's full name (Text/String)
specialization	Area of medical expertise (Categorical)
contact_number	Phone number of the doctor (Text/String)

billing

Attribute	Description
billing_id	Unique ID for each billing record (Categorical)
appointment_id	Reference to the appointment (Categorical)
service_charges	Charges for the medical service (Numerical)
medicine_charges	Cost of prescribed medicines (Numerical)
total_amount	Total bill amount (sum of above charges) (Numerical)

treatments

Attribute	Description
treatment_id	Unique ID for each treatment entry (Categorical)
appointment_id	Reference to the appointment (Categorical)
treatment_type	Type of treatment (e.g., Surgery, Therapy, Diagnosis) (Categorical)
treatment_notes	Additional notes about the treatment (Text/String)

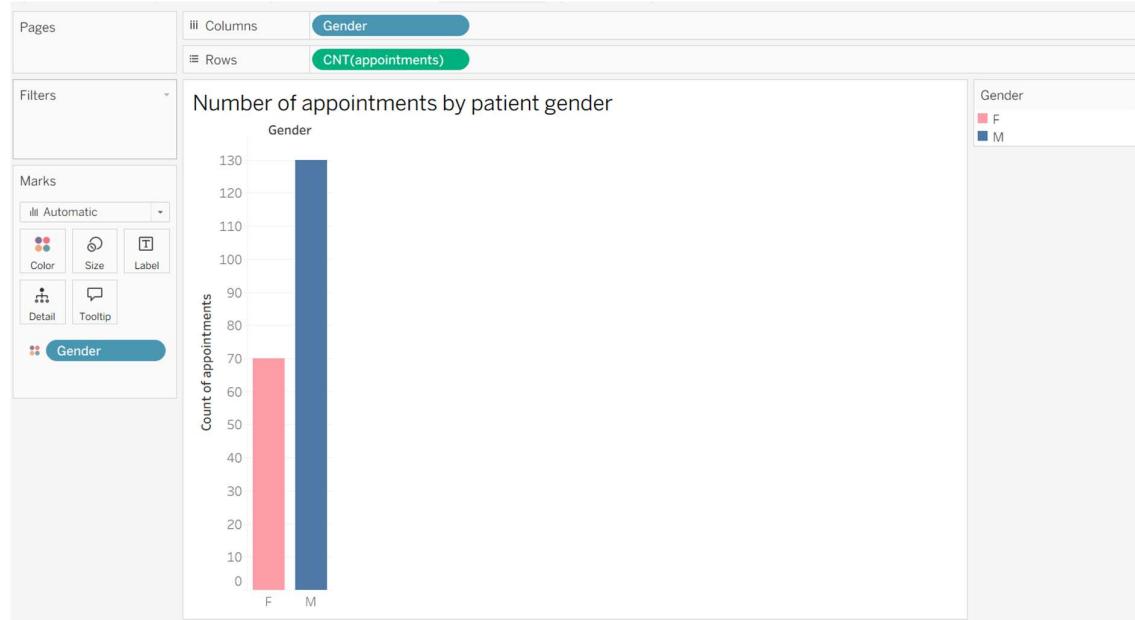
◆ Data Joining

Join 1: appointments + patients

- Joined on: patient_id
- Insight: Number of appointments by patient gender
- Chart Used: Bar Chart

The screenshot shows a data joining interface. On the left, under 'Connections', 'appointments.csv' is selected from 'Microsoft Excel'. Under 'Sheets', 'appointments' is listed. In the center, a tooltip for 'appointments' states it's made of 2 tables: 'appointments' and 'patients' connected by 'patient_id'. Below this, the 'appointments' table schema is shown with 18 fields and 200 rows. The columns include Appointment Id, Patient Id, Doctor Id, and Appointment Date. On the right, a legend for 'Gender' shows 'F' in pink and 'M' in blue.

Chart:



Join 2: appointments + doctors

- Joined on: doctor_id
 - Insight: Number of appointments per doctor
 - Chart Used: Horizontal Bar Chart

File Data Window Help

Tableau Desktop Public Edition Buy Tableau

Connections Add

appointments.csv Microsoft Excel

Sheets

- appointments
- billing
- doctors
- patients

New Union

New Table Extension

appointments+ (appointments.csv)

1 Alert

appointments1 is made of 2 tables.

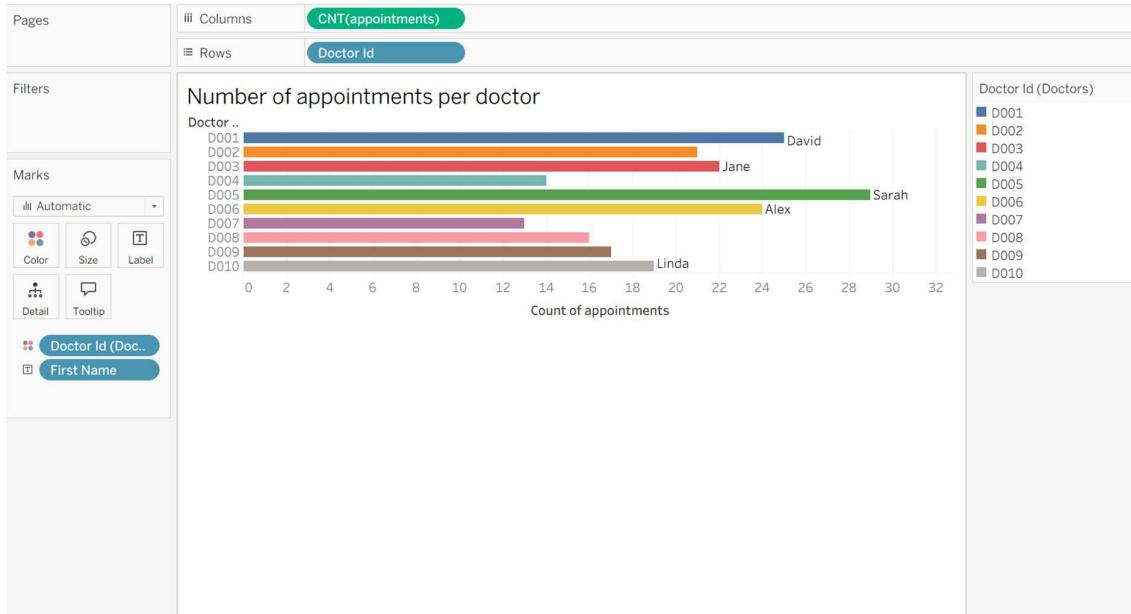
appointments1 doctors2

appointments1 15 fields 200 rows 100 rows

Name	Field Name	Type	Physic...	Rem...
Appointment Id (Appointm...	Abc	Appointment Id (Appointm...	appointment...	appointment...

Abc appointments1	Abc appointments1	Abc appointments1
Appointment Id (Appointm...	Patient Id (Appointments1	Doctor Id (Appointments1
A001	P034	D009
A002	P032	D004
A003	P048	D004
A004	P025	D006
A005	P040	D00?

Chart:



◆ Data Blending

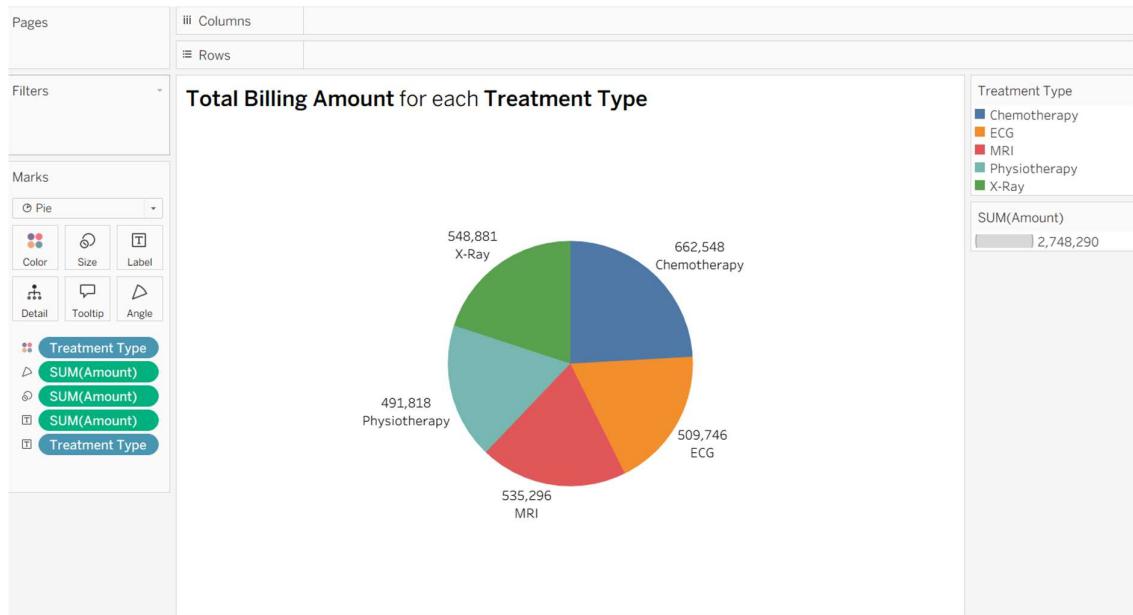
Blend: appointments (Primary) + billing (Secondary)

- **Linked using: appointment_id**
- **Insight: Total revenue generated per treatment type**
- **Chart Used: Pie Chart**

The screenshot shows the Tableau interface with the following details:

- Connections:** 'appointments.csv' (Microsoft Excel) is connected.
- Sheets:** 'appointments', 'billing', 'doctors', 'patients', 'New Union', and 'New Table Extension' are listed.
- Relationships:** A relationship is drawn between 'appointments' and 'billing' using the 'Patient Id' field.
- Preview:** A table view shows the joined data with columns: 'Bill Id', 'Patient Id (Billing)', 'Treatment Id', 'Bill Date', and 'Amount'.

Chart:



---X---