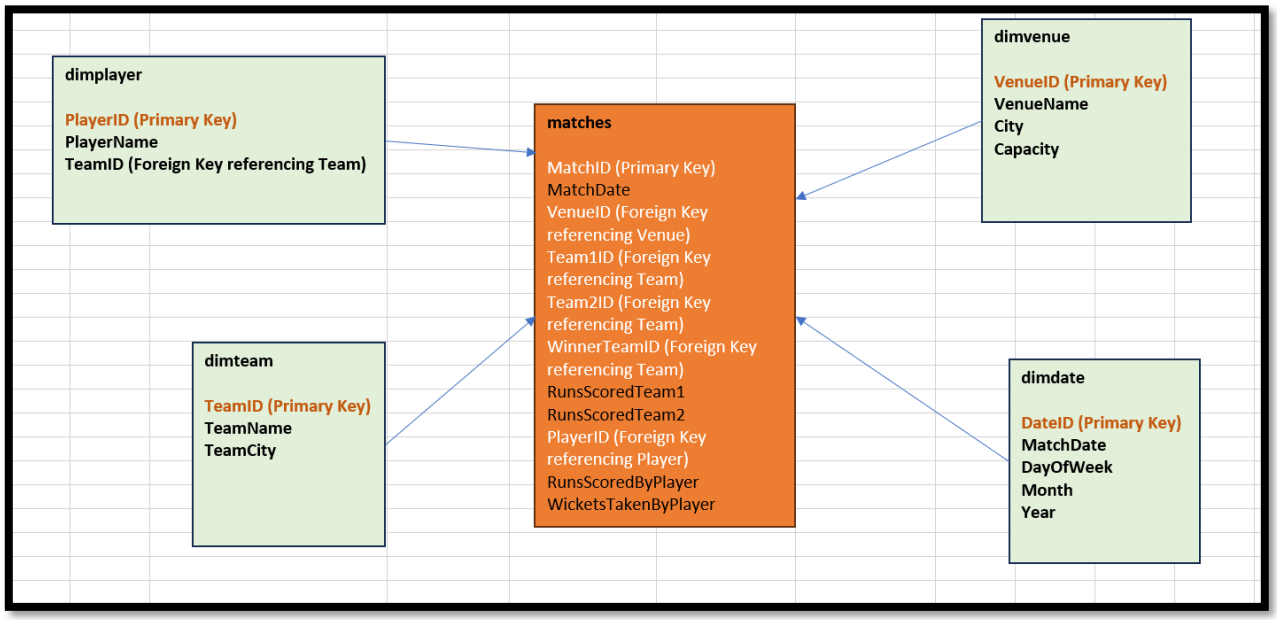
Design Data Warehouses for Given Below Products:

Note: While designing any Data Warehouse make sure to cover given below points.

- Design Fact & Dimension tables
- Create meaningful Primary & Foreign keys
- Try to follow Star/SnowFlake Schema Design
- Try to write few SQL queries to generate insightful business metrics (This is the critical point because you need to understand the Data & Business both)

1. Design a Data Warehouse for IPL Cricket Tournament (Asked in Flipkart Interview for Senior Data Engineer role)

businessprocess	grain	dimensions	facts
team and player performance.	one IPL match	team	matches
		player	
		venue	
		date	



```sql
-- Top Performing Teams - Total Runs Scored:
SELECT
    DimTeam.TeamID,
    DimTeam.TeamName,
    SUM(COALESCE(FactIPLMatchPerformance.RunsScoredTeam1, 0) + COALESCE(FactIPLMatchPerformance.RunsScoredTeam2, 0)) AS TotalRuns
FROM
    DimTeam
LEFT JOIN
    FactIPLMatchPerformance ON DimTeam.TeamID IN (FactIPLMatchPerformance.Team1ID, FactIPLMatchPerformance.Team2ID)
GROUP BY
    DimTeam.TeamID, DimTeam.TeamName
ORDER BY
    TotalRuns DESC
LIMIT 5;
```

Please click here for more information.

Output Messages Notifications

teamid [PK] integer	teamname character varying (255)	totalruns bigint
2	Team B	710
1	Team A	710

```sql
-- Player of the Tournament - Maximum Runs Scored:
SELECT
    DimPlayer.PlayerID,
    DimPlayer.PlayerName,
    SUM(COALESCE(FactIPLMatchPerformance.RunsScoredByPlayer, 0)) AS TotalRuns
FROM
    DimPlayer
LEFT JOIN
    FactIPLMatchPerformance ON DimPlayer.PlayerID = FactIPLMatchPerformance.PlayerID
GROUP BY
    DimPlayer.PlayerID, DimPlayer.PlayerName
ORDER BY
    TotalRuns DESC
LIMIT 1;
```

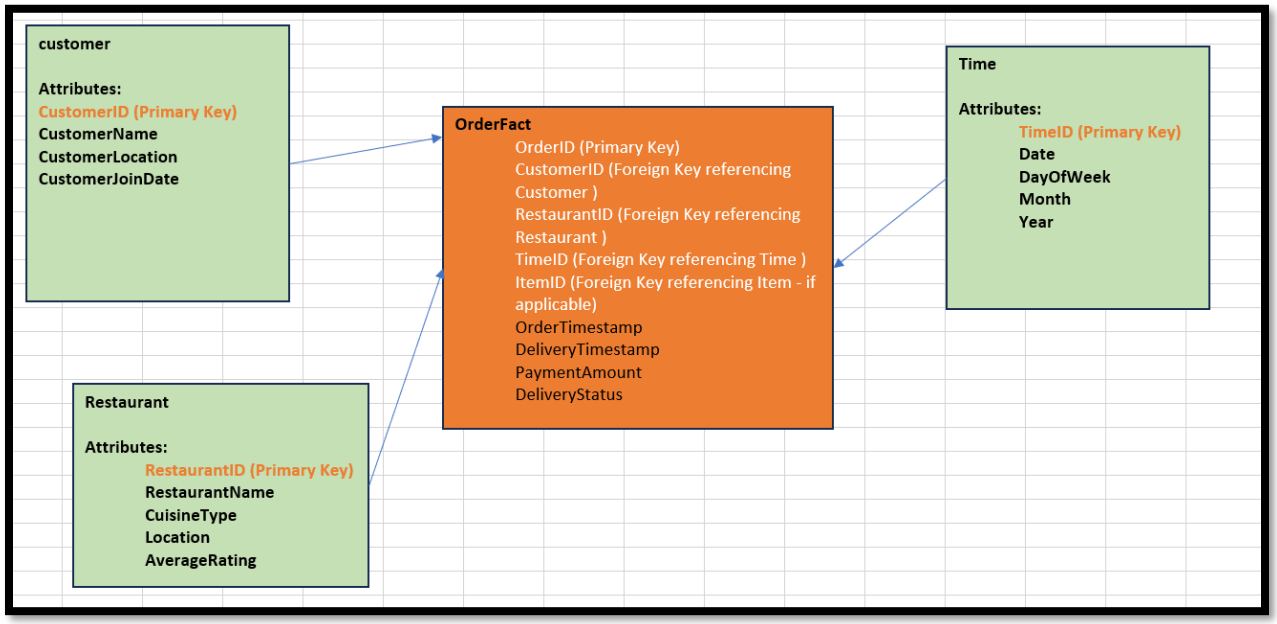
Please click he

Output Messages Notifications

playerid [PK] integer	playername character varying (255)	totalruns bigint
102	Player 2	80

2. **Design a Data Warehouse for Food delivery app like Swiggy, Zomato (Asked in Grab for Data Engineer role)**

Column1	Column2	Column3	Column4
business process	grain	dimension	fact
order management	single order made by customer	customer	
restaurant management		restaurant	
customer management		time	
delivery logistics			
menu management			
payment processing			
ratind and review			




```
57  -- total number of orders
58  SELECT COUNT(OrderID) AS TotalOrders
59  FROM OrderFact;
60
61
```

Data Output | Messages | Notifications

	totalorders bigint
1	2

```sql
-- most popular cuisine type
SELECT CuisineType, COUNT(RestaurantID) AS RestaurantCount
FROM Restaurant
GROUP BY CuisineType
ORDER BY RestaurantCount DESC;
```

Data Output | Messages | Notifications

	cuisinetype character varying (255)	restaurantcount bigint
1	Japanese	1
2	Italian	1

```sql
--average delivery time
SELECT
    r.RestaurantID,
    r.RestaurantName,
    AVG(of.DeliveryTimestamp - of.OrderTimestamp) AS AvgDeliveryTime
FROM Restaurant r
JOIN OrderFact of ON r.RestaurantID = of.RestaurantID
GROUP BY r.RestaurantID, r.RestaurantName;
```

Data Output | Messages | Notifications

	restaurantid [PK] integer	restaurantname character varying (255)	avgdeliverytime interval
1	101	Best Pizza	00:30:00
2	102	Tasty Sushi	00:45:00

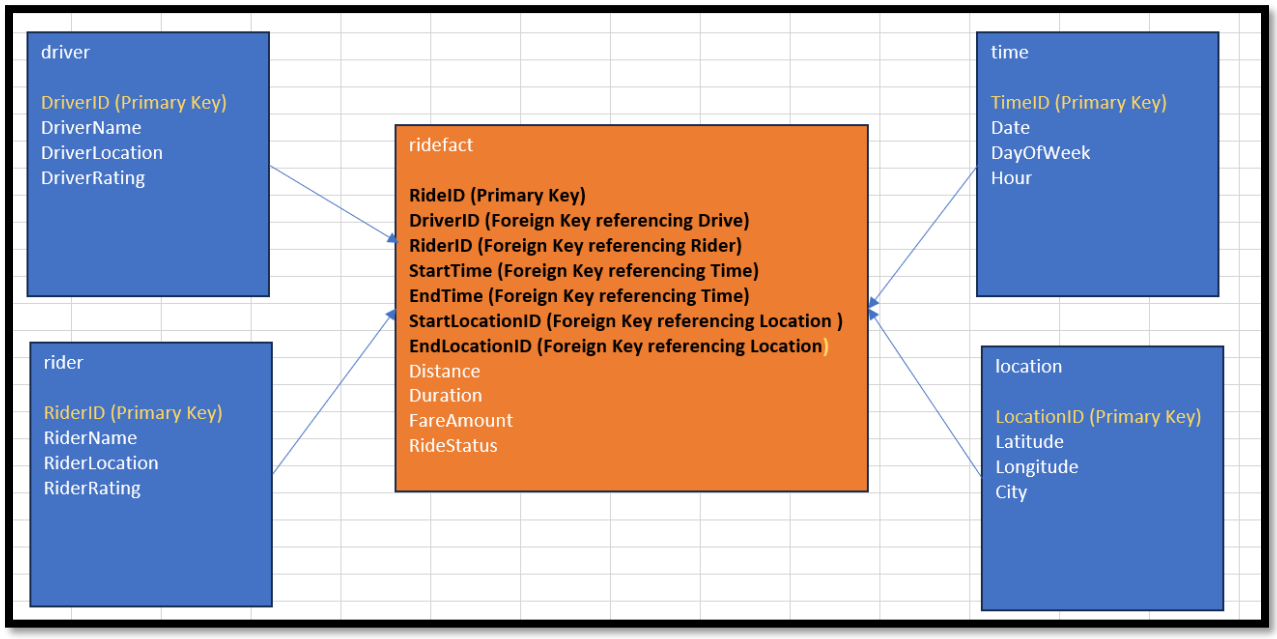
```sql
-- average payment amount per order
SELECT AVG(PaymentAmount) AS AvgPaymentAmount
FROM OrderFact;
```

Data Output | Messages | Notifications

	avgpaymentamount numeric
1	27.7500000000000000

3. **Design a Data Warehouse for cab ride service like Uber, Lyft (Asked in Google for Data Engineer role)**

Column1	Column2	Column3	Column4
business process	grain	dimensions	fact
ride booking		driver	
payment processing		rider	
driver and rider management		time	
		location	



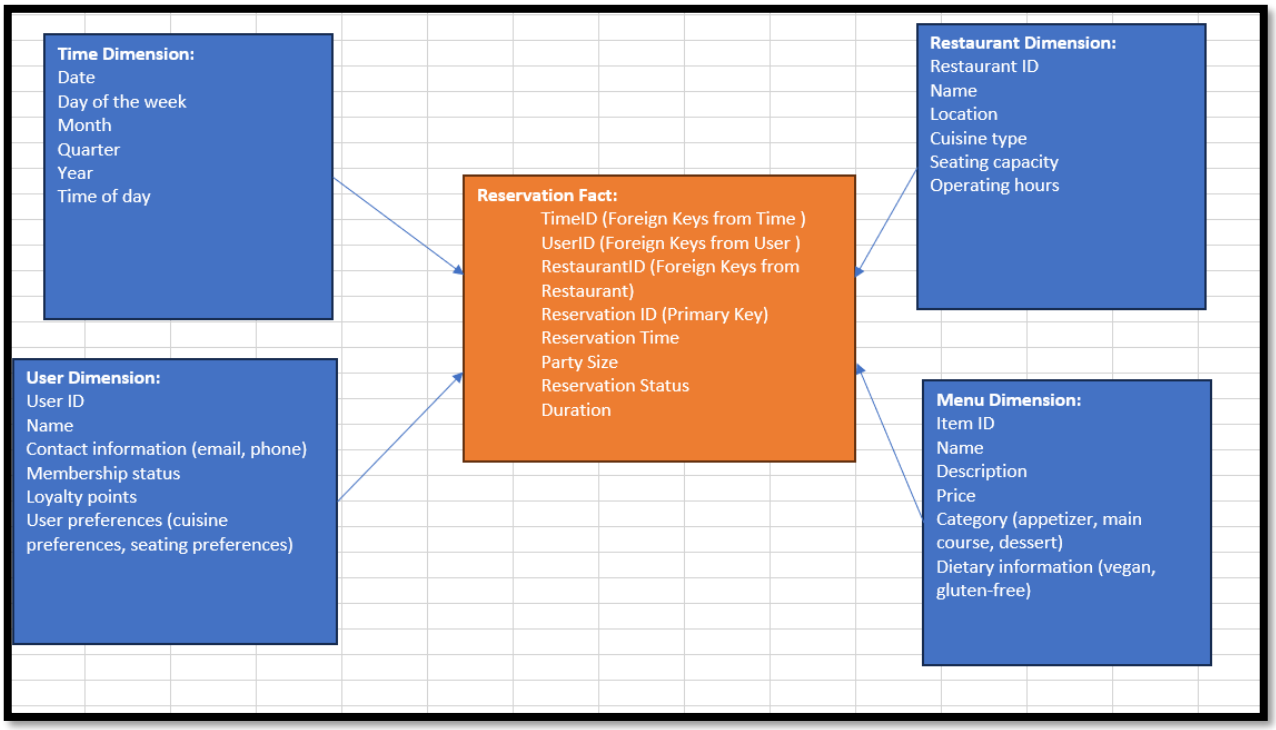

```sql
-- Average Duration of Rides by Driver Rating:
SELECT
    d.DriverRating,
    AVG(rf.Duration) AS AvgDuration
FROM DriverDimension d
JOIN RideFact rf ON d.DriverID = rf.DriverID
GROUP BY d.DriverRating;
```

Output Messages Notifications

driverrating numeric (3,2)	avgduration numeric
4.60	25.0000000000000000
4.80	30.0000000000000000

4. Design a Data Warehouse for Restaurant table booking app like Dineout (Asked in McKinsey for Consultant Data Engineer role)

Column1	Column2	Column3	Column4
business process	grain	dimensions	fact
user reservation	individual reservation and UI	time	
restaurant management		user	
user interaction		restaurant	
		menu	




```
74  -- Query 1: Get reservations made on a specific date
75  SELECT *
76  FROM ReservationFact
77  JOIN TimeDimension ON ReservationFact.TimeID = TimeDimension.TimeID
78  WHERE TimeDimension.Date = '2024-02-17';
79
```

Data Output | Messages | Notifications

	reservationid integer	timeid integer	userid integer	restaurantid integer	partysize integer	reservationstatus character varying (20)	duration integer	reservationtime timestamp without time zone	timeid integer	date date	dayofweek character varying (10)
1	101	1	1	1	4	Confirmed	90	2024-02-17 18:30:00	1	2024-02-17	Wednesday

```sql
-- Query 3: Find the average party size for reservations made by Gold members
SELECT AVG(ReservationFact.PartySize) AS AvgPartySize
FROM ReservationFact
JOIN UserDimension ON ReservationFact.UserID = UserDimension.UserID
WHERE UserDimension.MembershipStatus = 'Gold';
```

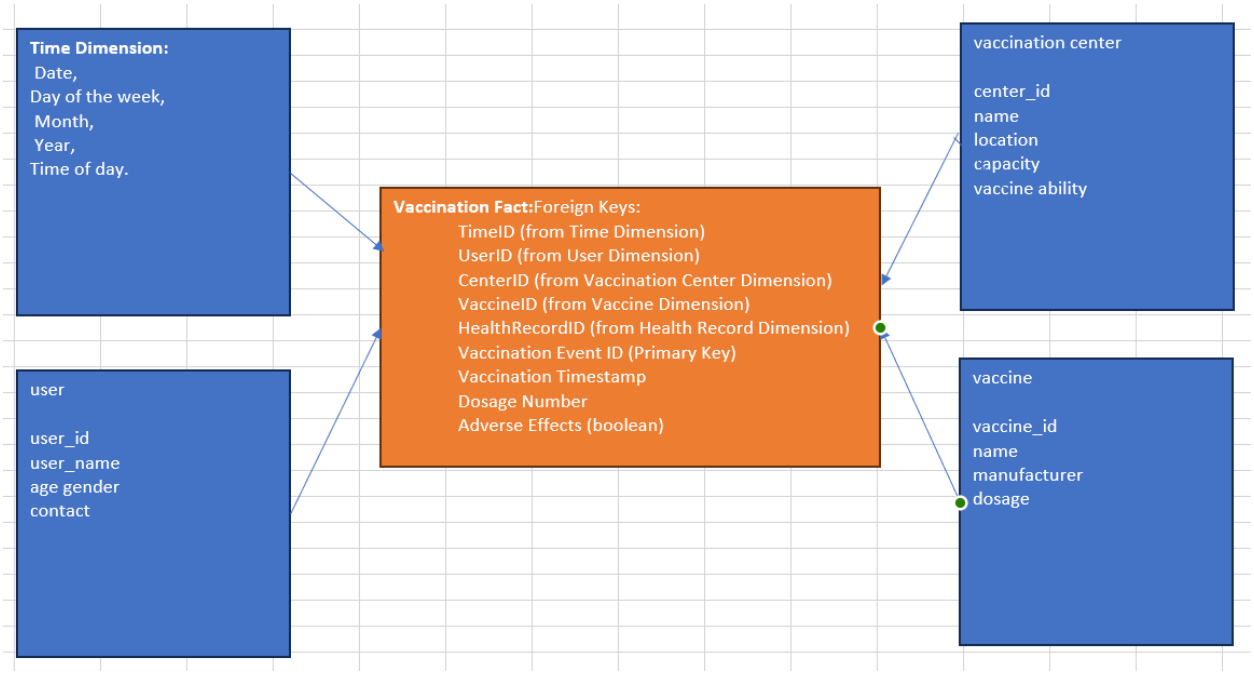
Output Messages Notifications

avgpartysize numeric
4.0000000000000000

5. **Design a Data Warehouse for Covid Vaccination Application (Asked in Livsapce for Data Engineer role)**

Column1	Column2	Column3	Column4
business process	grain	dimensions	fact
user registration	single vaccination event	time	
vaccination centers		user	
appointment		vaccination center	
vaccination		vaccine	
health records			

SHRUTI ZALAVADIYA




```sql
-- Query 3: Get the count of vaccination events by vaccine type
SELECT VaccineDimension.Name, COUNT(*) AS VaccinationCount
FROM VaccinationFact
JOIN VaccineDimension ON VaccinationFact.VaccineID = VaccineDimension.VaccineID
GROUP BY VaccineDimension.Name;
```

Output Messages Notifications

name character varying (50)	vaccinationcount bigint
COVID-Vax2	1
COVID-Vax1	1