= 1st problem =

Assume we have the following table and columns

Table: UnitsSold

   ProductID

   Quantity

   Date

  ForecastOrActual – can be “A” for actual and “B” for forecast.

Sample Data would look like

|  |  |  |  |
| --- | --- | --- | --- |
| **ProductID** | **Quantity** | **Date** | **ForecastOrActual** |
| Widget | 100 | 1-June-2015 | F |
| Widget | 120 | 1-June-2015 | A |
| Gizmo | 90 | 1-June-2015 | F |
| Gizmo | 80 | 1-June-2015 | A |
| Doohickey | 200 | 1-June-2015 | A |

 Write an SQL query which returns a single metric indicating how well actuals met forecast for a given week.  What assumptions did you make?

Required output: *single metric* indicating how well actuals met forecast for a *given week.*

Assume, the Forecast and Actual Quantities of a ProductID are placed side-by-side by massaging the input data into a new table called UnitsSoldPlus. A sample is shown below.

|  |  |  |  |
| --- | --- | --- | --- |
| **ProductID** | **ForecastQuantity** | **ActualQuantity** | **Date** |
| Widget | 100 | 120 | 1-June-2015 |
| Gizmo | 90 | 80 | 1-June-2015 |
| Doohickey | 200 | 200 | 1-June-2015 |

With this format, it is easy to calculate Mean Absolute Error or MAE, which is a “quantity used to measure how close forecasts or predictions are to the eventual outcomes.” [1]

SELECT r.sumerror/r.numOfRecords AS MAE

FROM (

SELECT COUNT(\*) AS numOfRecords,

SUM(ABSOLUTE(ForecastQuantity – ActualQuantity)) AS sumerror

FROM UnitsSoldPlus

WHERE DATE BETWEEN ‘$startDate’ AND ‘$endDate’

) AS r;

$startDate and $endDate represent start and end date of a given week, for example, 1-June-2015 and 7-June-2015.

Assumptions,

1. $startDate and $endDate represent a given week boundary.
2. Both forecast and actual quantity values are present for a given ProductID.

[1] <https://en.wikipedia.org/wiki/Mean_absolute_error>

= 2nd problem =

Given Truck A arrives at a random time between 9am and 11am, and Truck B arrives at a random time between10am and 12am.  What are the Odds that Truck A arrives before Truck B?  Why

The probability that Truck A arrives before Truck B is 0.5625 .

Method:

1. Divide the time for Truck A in two windows, namely 9-to-10 and 10-to-11. The probability of A randomly arriving in the 9-to-10 window is ½ -- similar to a coin toss.
2. Similarly, divide the time for Truck B in two windows, namely 10-to-11 and 11-to-12. The probability of B randomly arriving in the 11-to-12 window is ½.
3. Now, subdivide the time in 10-to-11 in two parts:
   1. 10 + x minutes, named as early part, where x < 60 min
   2. the probability that Truck A arrives in the early part at a random time = the probability that Truck A arrives in the 10-to-11 window (i.e. ½) AND the probability that Truck A happens to be in 10+x window (i.e. ½) = ½ x ½ = ¼
   3. the probability that Truck B NOT arrives in the early window = the probability that Truck B arrives in the 10-to-11 window (i.e. ½) AND the probability that Truck B is NOT in the early window (i.e. ½) = ½ x ½ = ¼
4. The probability of Truck A arriving in 9-to-10 OR 10+x window = I + b = ½ + ¼ = 0.75
5. The probability of Truck B arriving in 11-to-12 OR NOT(10+x) window = II + c = ½ + ¼ = 0.75
6. The probability of Track A arriving before Truck B = IV x V = 0.75 x 0.75 = 0.5625