**NURTITION**

**For this table I did two approaches**

1. taking patients having basic nutrition (by comparing the columns in the query) : if it's greater than 3, they will come under 'yes' or 'no'.

with cte as (

select case\_id,breakfast\_meal,lunch\_meal,meal\_dinner,vegetables,bean,fruits,prepartum\_maternal\_weight,

case when breakfast\_meal + lunch\_meal + meal\_dinner + vegetables + bean + fruits >= 3 THEN 1

else 0 end as basic\_nurtition

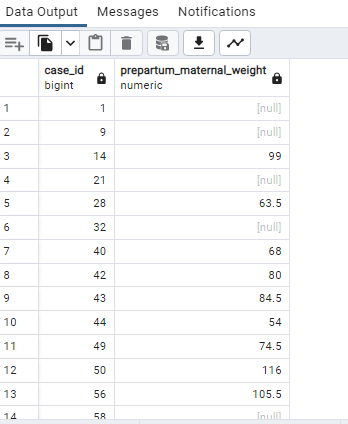
from public.maternal\_patient\_data )

select case\_id,prepartum\_maternal\_weight

from cte

where basic\_nurtition = 0

order by case\_id



so, as per this logic, 72 those who didn't meet basic nutrition. I compared with different columns, weights, and glucose but didn't get enough data. If needed, we can drop.

**Approach 2**: I added cookies and pasta (if patients took both, then they will come under high-intake calories). Initial comparison I did with glucose and a few.

with cte as (

select case\_id,pasta,cookies,newborn\_weight,prepartum\_maternal\_weight,preeclampsia\_record\_pregnancy,

case when pasta+cookies >=2 then 1

else 0

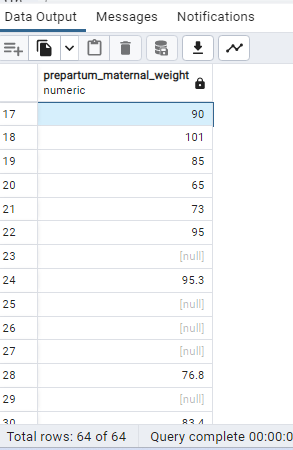
end as regular\_high\_calorie\_intake

from public.maternal\_patient\_data)

select prepartum\_maternal\_weight

from cte

where regular\_high\_calorie\_intake = 1;



where here I got 64 patients both using cookies and pasta regularly. If we see with prepartum weight (17 patients' weight is 80 and above), might be this will be helpful in further analysis."

2(a) if we compare patients with 0 in basic nutrition and 1 in high intake calories, I'm getting 16 patients (6) in above weight.

2(b) if using 'or' (we're getting patients above 80 weight, 30).

This is my initial research on nutrition. I need your suggestions. Will it be helpful in our analysis? Do we need to think in a different way (by adding and removing certain columns)?if we go by this approach, we will have basic\_nurtition\_column and high\_calorie\_column for nutrition, and we can drop the remaining.

**Hypertension**

right\_systolic\_blood\_pressure

left\_systolic\_blood\_pressure Mean\_systolic\_bp

left\_diastolic\_blood\_pressure

right\_diastolic\_blood\_pressure mean\_diastolic\_bp

**drop**

We can drop the records in the 'hypertension\_past\_treatment' column where the value is 'yes,' and details or clarity are not available for these 3 patients. These patients are already included in another column, 'hypertension\_past\_reported,' where their entries exist

**labor\_admission\_day / new\_born (LABOR and delivery/ admission and DELIVERY)**

Set the delivery mode based on numeric values ranging from 0 to 12 in the c-section column. To simplify interpretation, we can categorize them into 0, 1, or 2. This approach helps avoid further confusion. Additionally, in cases where the cesarean section column contains numeric values, comparing them with the delivery value provides a clear idea of where they should be placed.

The **Delivery\_mode** column is coded as follows:

* 0: Vaginal
* 1: Caesarian
* 2: Others (includes any reason, missed follow-up, not included in the study)

For example, in the case of case\_id 146, there is data for newborn weight, but related columns, including Delivery\_mode, have no values. It is noted that the individual did follow-up later, and they might not be included in the study for various reasons.

**mothers\_hospital\_stay** (The approach taken to replace null values in this column is as follows)

The **mothers\_hospital\_stay** column has been updated for null values, specifically 'same\_day\_discharge' and 'missed\_follow\_up.' This update is based on consideration of values such as newborn weight, Apgar scores at 1 hour, and Apgar scores at 5 hours.

* Type conversion need to done for all columns in this table

**PREVIOUS\_PREGANCY**

We can consider deleting all columns related to **past\_newborn\_weight\_1** through **past\_newborn\_weight\_4** because

**Findings:**

In the "Past Pregnancy" column, there are 9 null values. However, the related "Weight" column has values, and currently, we can consider them valid. To address this, we can replace the null values in the "Past Pregnancies Number" column by adding to a respective column indicating the availability of data.

**Next Steps:**

If we move to the next case, focusing on those who had exactly 1 past pregnancy, we observe that there are 99 columns. Unfortunately, we have data for only a few columns, less than 10. In this scenario, we can make the assumption that:

Yes, they have had one past pregnancy, but they didn't deliver here.

Or, we don't have enough data for the records.

Considering the lack of data, there is a suggestion to **drop** the "Weight" columns and use the "Past Pregnancy" column to understand the patient's previous history by either replacing null values or keeping them as they are.

**Gestational Age for Past Newborns:**

The column "Gestational Age Past Newborn" classifies gestational age into two categories:

0: Preterm

1: Full term

**Considerations for Preterm Babies:**

It's noteworthy that the weight of all preterm babies is not constant, and when we consider the current newborn weight, it is observed to be different. Specifically, the current newborn weight is not low; it varies.

**Decision to Drop the Column:**

In light of the differences observed in the weights of preterm babies and the consideration of current newborn weight, it is suggested that we can drop the "Gestational Age Past Newborn" column.

**DISEASE**

In the cleaning steps, we need to update the values in the columns **chronic\_diseases**, **disease\_diagnose\_during\_pregnancy**, and **treatment\_disease\_pregnancy** as follows:

1. Change 'null' to '0'.
2. Change 'not\_applicable' to '0'.
3. Change 'no\_answer' to '0'.

Additionally, during the cleaning process, we observed anomalies in row 4 (case\_id 120) where numeric values exist in the respective columns. Furthermore, case\_id 120 has 'thb' in a column, and we are uncertain about its definition. However, we cannot modify this value as the patient is taking the medicine 'ac\_valproico,' which is not advisable during pregnancy.

Moreover, row 244 has a different value that seems to have no meaningful definition.

**Gestational\_column**

For case\_id 47 gestational\_age\_birth less than age\_inclusion (have to change in cleaning ?)

**WEIGHT\_COLUMN**

We can update the **current\_maternal\_weight\_1st\_tri**, **current\_maternal\_weight\_2nd\_tri**, and **current\_maternal\_weight\_3rd\_tri** columns based on certain conditions related to **gestational\_age\_at\_inclusion** and **maternal\_weight\_at\_inclusion**. The updated values are determined as follows:

* For **current\_maternal\_weight\_1st\_tri**:
  + If gestational\_age\_at\_inclusion is between 0 and 13, and maternal\_weight\_at\_inclusion is greater than the current value or the current value is NULL, set it to maternal\_weight\_at\_inclusion.
* For **current\_maternal\_weight\_2nd\_tri:**
  + If gestational\_age\_at\_inclusion is between 13 and 27, and maternal\_weight\_at\_inclusion is greater than the current value or the current value is NULL, set it to maternal\_weight\_at\_inclusion.
* For **current\_maternal\_weight\_3rd\_tri**:
  + If gestational\_age\_at\_inclusion is between 27 and 40, and maternal\_weight\_at\_inclusion is greater than the current value or the current value is NULL, set it to maternal\_weight\_at\_inclusion.

**DRop**

We can then consider dropping the maternal\_weight\_at\_inclusion column as its values have been incorporated into the updated columns."

Comparsion columns

alcohol\_use,alcohol\_preference,alcohol\_quantity\_milliliters

**When alcohol\_use is equal to 1, and the other two columns have null or are not applicable, change them to 'no answer’**

**When alcohol\_use is equal to 0, and the other two columns have null or are not applicable, change them to 'not used**.

Comparsion columns

tobacco\_use\_in\_months,tobacco\_use,tobacco\_quantity\_by\_day

**If tobacco\_use is zero, and tobacco\_use\_in\_months is null or not\_applicable, then null or not\_applicable values in tobacco\_quantity\_by\_day are changed to 'not\_used.' If tobacco\_use is equal to 1 and tobacco\_use\_in\_months has a value, then tobacco\_quantity\_by is changed to 'no\_answer.**

**Drugs\_prefrence**

**'Not applicable' value changed to '0' when drugs\_prefernce is Zero and drugs\_years\_use is 'Not applicable'**

**Null value changed to 'No\_anwer' drugs\_prefernce !=0 and drugs\_years\_use is 'Not applicable'**

<https://nationaltasc.org/types-of-alcohol/>

future ref