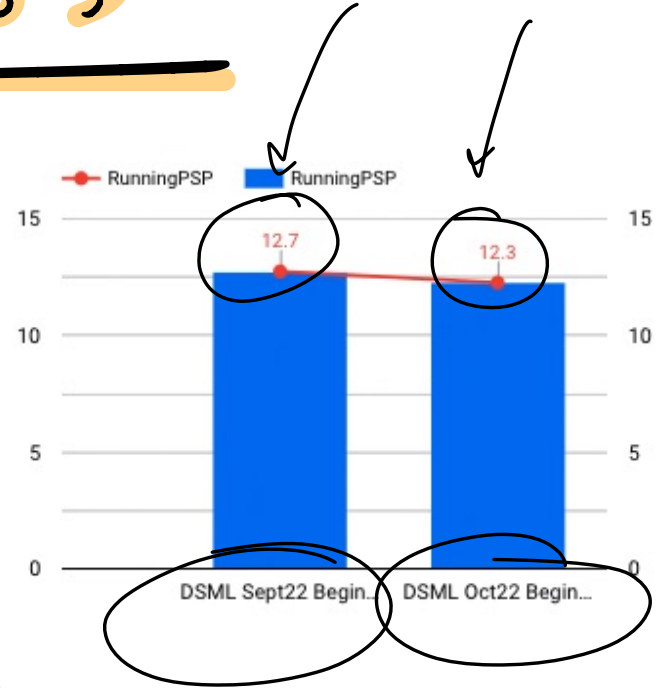
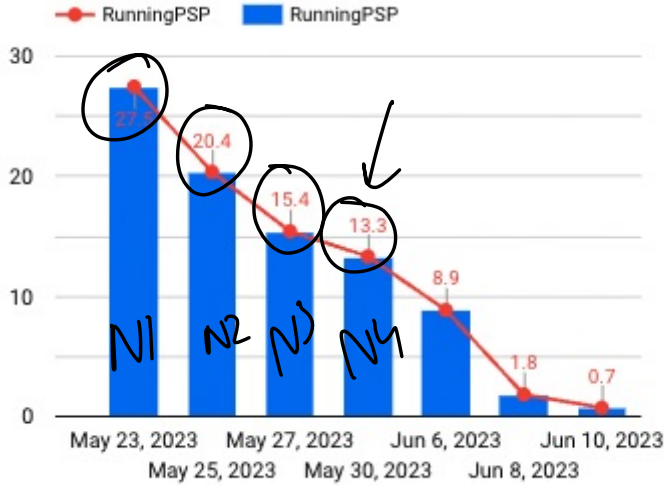
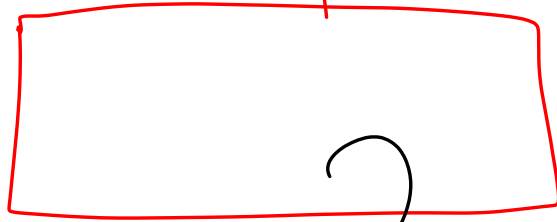


PANDAS-5

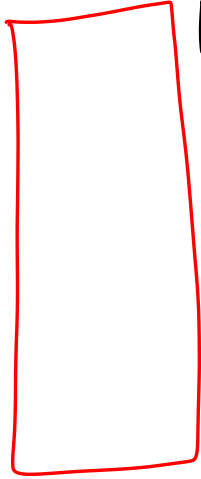


Please Solve the Question:

wide format



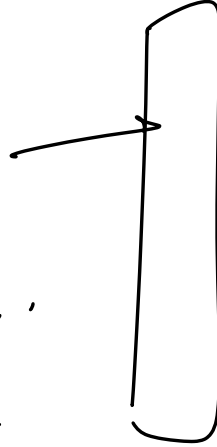
Date	DIN	Press	Temp	Time
1	1	0	0	1:30
1	1	0	0	2:30
1	1	0	0	3:30



long format


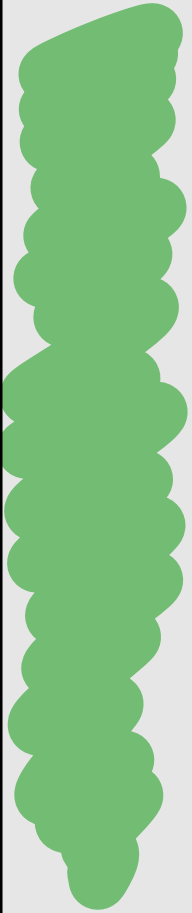
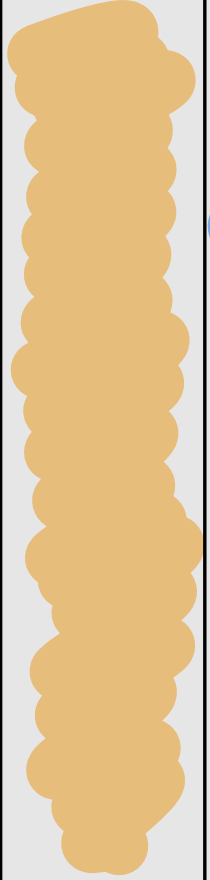


wide to long.

pd.melt,



long → wide
pd.pivot

DATE DRUG Name PARAM 1:30 2:30 3:30 - - - 12:00

		T P T P T P				
---	--	----------------------------	---	--	---	--

DATE	DN	PARAM	Time	Value
		T P	1.30	
		T P		
		T P		
		T P		
		T P		
		T P	2.30	
		T P		
		T P		
		T P		
		T P	2.30	

Handle Missing Values

100 datapoints

$x\%$ of missing value

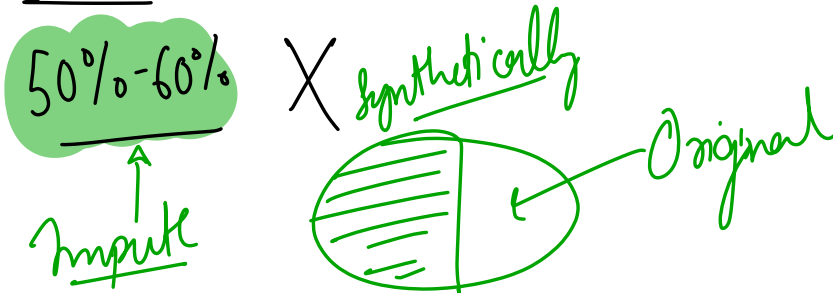
→ DROP

→ Impute →
Continuous → Mean
Categorical → Median
→ Mode

5% DROP

10% DROP/IMPUTE

20-30% IMPUTE



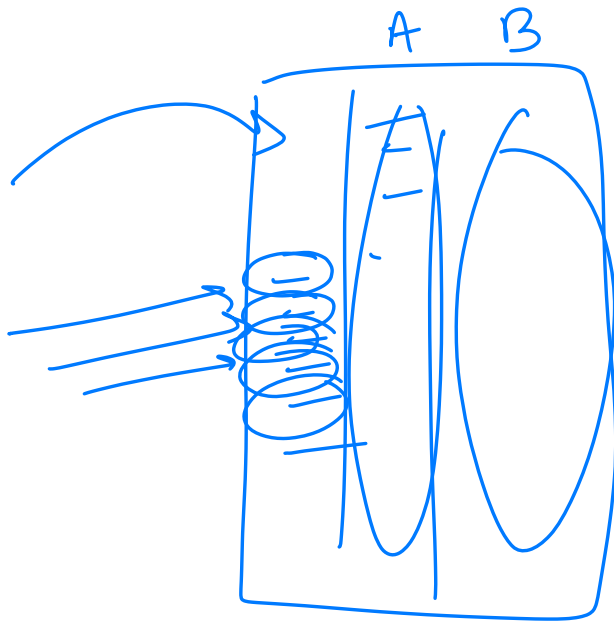
PAYPAL Transaction data X

STRIPE X

AMAZON 5% ✓

HEALTHCARE ← Impute

NETFLIX 10%



price	ID	Product ID
20	P1	X1
20	P2	X2
36	P3	X3
40	P1	X3
50	P2	X2
20	P3	X1