

Clarification questions:

- What is the drug form of staquili ? Or Is it that the product consists of different types of drugs which are mentioned?
- Which metric are we tracking for considering top payer name?
- In Q6. what is the number of top payers required (Top1, Top 10, Top 100) ?
- What factors/metrics are required to track for understanding more imp payer channels?

Q1. What types of Data checks you'll apply to the provided data to ensure the integrity and completeness of the given data tables?

SOLUTION:

1) Data Exploration Checks

- By exploring the data we found I found that **Service data** in the **Claims** is not in the correct format.
- Hence converted the complete data in the correct(Date) format.

	A	B	C	D	E	F	G	H	I	J	K
1	Claim ID	Patient ID	NDC	Service Date	Quantity	Plan ID	Data Source	Physician ID	Diagnosis Code		
30	10308099445001689940	102686832	50090023600	43324	454	6811811	RX	6800371	-		
31	10577197139300161894	102890080	-	42845	0	6796601	MX	6796464	L20		
32	10577197139300824351	102890080	-	42845	0	6812400	MX	10605421	L20		
33	10047190318900863264	102914824	-	42782	0	7964107	MX	8028543	L20		
34	10490194259811427510	102914824	21695050180	43147	45	16707226	RX	6810428	-		
35	165165091259	103014611	-	43116	0	12033139	MX	6796449	L20		
36	165165091257	103014611	-	43116	0	6800758	MX	16318502	L20		
37	166423757286	103014611	51672201309	43131	90	6796762	RX	6805430	-		
38	10047092138800447779	103014611	40085002880	43193	60	6800149	RX	6430426	-		
39	10047093825000441915	104011131	-	42859	0	6806007	MX	6797031	L20		
40	10047096703500480267	104011131	51672127001	42912	45	6796496	RX	6813833	-		
41	10015134668500054119	104876896	-	43725	0	18576550	MX	6810785	L20		
42	10015134668500054121	104876896	-	43725	0	6799544	MX	6812054	L20		
43	10015142503000055696	104876896	54569227500	43802	90	6812731	RX	10453104	-		
44	10577192245801211621	104963691	-	42799	0	6810818	MX	6811304	L20		
45	10606194899803464146	104963691	50090023700	43189	454	6809510	RX	6791548	-		
46	10725127924000323985	105126596	-	43664	0	6805408	MX	6813586	L20		
47	10725129042600122142	105126596	50090327600	43675	45	6812441	RX	6800872	-		
48	161503836502	105154384	-	43441	0	15154081	MX	6812506	L20		
49	507964739589	105154384	54569420000	43500	180	6807892	RX	6417440	-		
50	511808157690	105619000	-	43318	0	6801015	MX	6810425	L20		
51	511829716346	105619000	11822372390	43321	45	6798863	RX	6806891	-		
52	165216933634	105704284	-	42752	0	6798226	MX	6800218	L20		
53	10490088668704775382	105704284	47781053226	43118	180	6792417	RX	6798909	-		
54	501382556263	105728145	-	42892	0	6811557	MX	6810309	L20		
55	507650262655	105728145	50090164700	43468	30	6789984	RX	6813705	-		
56	509333041961	105728145	55724021123	43700	90	6811960	RX	6789687	-		

-Formatting is done to convert it into date format

Format Painter | Font | Alignment | Number | Styles

Undo | Clipboard | Font | Alignment | Number | Styles

A19322 | 10490098215104551525

	Claim ID	Patient ID	NDC	Service Date	Quantity	Plan ID	Data Source	Physician ID	Diagnosis Code	J	K	L	M
19321	165097445215	2298912933	45802039001	12-01-2018	180	6797065	RX	6431899	-				
19322	10490098215104551525	2299806104	-	19-07-2017									
19323	501168695802	2300364021	-	16-05-2017									
19324	10420086207602456455	2301931396	-	15-11-2017									
19325	10490101917800125217	2320008403	-	13-09-2017									
19326	10236101638103807353	2320236890	-	12-09-2017									
19327	10193086541600737189	2320276597	-	01-12-2017									
19328	10490191069003184165	2320284747	-	23-01-2017									
19329	500055185757	2320284747	52959029915	27-01-2018	4								
19330	10490195931002728943	2320368243	-	15-01-2017									
19331	165166651804	2320368243	51672418803	16-01-2018	6								
19332	10419090767701590672	2320883092	-	24-02-2017									
19333	10606193663706177600	2320883092	50428030208	18-03-2018	3								
19334	10253097160500105694	2321945719	-	27-06-2017									
19335	10420099718003342601	2324835594	-	29-07-2017									
19336	10420100202102863094	2324914080	-	22-08-2017									
19337	10490093408402444244	2325015625	-	13-04-2017									
19338	10490091602700740766	2325034659	-	13-03-2017									
19339	10420100052101695975	2331201338	-	22-08-2017									
19340	10420087955200050909	2331326953	-	05-01-2017									
19341	164538856081	2331326953	54868487802	02-01-2018	1								
19342	176867070612	2332195280	-	11-11-2017									
19343	10420208381601951792	2332279000	-	27-07-2017									
19344	10420084898001208644	2333209376	-	01-11-2017									
19345	10490089277306214784	2333235063	-	27-01-2017									
19346	10832189919300417073	2333235063	11383018551	03-02-2018	4								
19347	10490092616903523059	2333342972	-	07-04-2017									
19348													

Format Cells

Number | Alignment | Font | Border | Fill | Protection

Category:

- General
- Number
- Currency
- Accounting
- Date
- Time
- Percentage
- Fraction
- Scientific
- Text
- Special
- Custom

Sample

10490098215104551525

Type:

- *14-03-2012
- *14 March 2012
- 14-03-2012
- 14-03-12
- 14-3-12
- 14.3.12
- 2012-03-14

Locale (location):

English (India)

Calendar type:

Gregorian

Date formats display date and time serial numbers as date values. Date formats that begin with an asterisk (*) respond to changes in regional date and time settings that are specified for the operating system. Formats without an asterisk are not affected by operating system settings.

OK Cancel

-Final

	Claim ID	Patient ID	NDC	Service Date	Quantity	Plan ID	Data Source	Physician ID	Diagnosis Code	J
38	10069138400100446501	100639327	55724021123	23-10-2019	45	6811228	RX	10579436	-	
39	166512394538	100640793	-	31-01-2017	0	8021238	MX	6812894	L20	
40	10419088517701589349	100640793	40986001936	10-01-2018	120	10261157	RX	10672374	-	
41	10606089381005632209	100640793	43598034190	01-02-2018	30	6800375	RX	6811249	-	
42	10740141892435133083	101036153	-	08-03-2018	0	7973285	MX	7843404	L20	
43	10740107241933773376	101036153	24470092060	22-09-2018	30	6791189	RX	6800776	-	
44	10047089465001025761	101092106	-	04-02-2018	0	8277306	MX	6805751	L20	
45	10047096841101932855	101092106	33261042501	28-06-2018	90	6800355	RX	6795846	-	
46	10047094378300248741	101363189	-	15-05-2017	0	6812192	MX	6797325	L20	
47	10047095244900226613	101363189	33261072701	31-05-2017	30	20879139	RX	6790285	-	
48	162241099797	101363189	54868158900	16-12-2017	45	10326795	RX	6793603	-	
49	162176325849	101363189	55724021111	15-12-2017	30	10081863	RX	11064409	-	
50	10047095949000212883	101363189	51672129501	13-06-2017	45	10117432	RX	15947885	-	
51	10047091394100759792	101363202	-	19-03-2017	0	5988069	MX	6796244	L20	
52	161383808293	101363202	50428445264	06-12-2017	180	6808750	RX	8055649	-	
53	161383808287	101363202	51672125903	06-12-2017	60	8749330	RX	6794746	-	
54	10420094110403569746	101578090	-	10-04-2018	0	6789636	MX	6799618	L20	
55	10420097740701323984	101578090	51672125906	15-06-2018	30	6812044	RX	6801613	-	
56	164980156702	101690733	-	14-01-2017	0	9872376	MX	6804611	L20	
57	10047098997000200742	101690733	16729042201	04-08-2018	45	6811972	RX	11949850	-	
58	164980156705	101690733	45802039001	14-01-2018	180	16447195	RX	11241714	-	
59	10047091975400862855	101789657	-	30-03-2018	0	6146928	MX	11101639	L20	
60	163352297852	101789657	55724021121	27-12-2018	60	7992058	RX	18944987	-	
61	10490093907408659749	101872171	-	26-04-2018	0	6802097	MX	6797871	L20	
62	500939605624	101872171	-	26-04-2018	0	6803364	MX	8010039	L20	
63	10493098970000151010	102091370	-	04-08-2017	0	6789501	MX	8913310	L20	
64	10493099094800195535	102091370	54868244602	07-08-2017	90	6811150	RX	15172806	-	
65	10493115461600222203	102091370	45802000402	19-03-2019	45	6806057	RX	7623751	-	

- In **Market Definition** table, the units for certain inputs is **MG/MG** and **MG/ACT**, while the majority of the data is in percentage. Hence we need to convert MG/MG and MG/ACT into percentage.

	A	B	C	D	E	F	G
1	NDC	Drug Name	Brand/Generic Fla	Drug Form	Drug Strengt	Drug Class	
495	10631009307	KENALOG SPRAY	Brand	AEROSOL, SOLN	0.147 MG/MG	Topical Corticosteroids	
496	10631009362	KENALOG SPRAY	Brand	AEROSOL, SOLN	0.147 MG/MG	Topical Corticosteroids	
667	21695091263	KENALOG SPRAY	Brand	AEROSOL, SOLN	0.147 MG/MG	Topical Corticosteroids	
884	45802004501	TRIAMCINOLONE ACTN	Generic	AEROSOL, SOLN	0.147 MG/MG	Topical Corticosteroids	
885	45802004502	TRIAMCINOLONE ACTN	Generic	AEROSOL, SOLN	0.147 MG/MG	Topical Corticosteroids	
983	49502053735	IMPEKLO	Brand	LOTION	0.15 MG/ACT	Topical Corticosteroids	
1547	54868325500	KENALOG SPRAY	Brand	AEROSOL, SOLN	0.147 MG/MG	Topical Corticosteroids	
1602							
1603							
1604							
1605							
1606							

2) Data Accuracy Test:

Quantity of the drug is negative for some entries in **Claims Data** which is not possible. Hence the data is wrong.

Deleted the rows from the data

[illegible]

3) Missing Data:

- The data of **Zip codes** of the regions is missing for some of the rows from **Specialty Description** table.

[illegible]

- The data of **Specialty Code** is missing for row in the **Specialty Description** table.

1	Physician ID	Last Name	First Name	City	Stat	ZIP Cod	Specialty Cod	Specialty Desc
781	6807459	CHRISTOPHER	FROST	SOMERSET	KY	42503		DERMATOLOGY
538								
539								
540								
541								
542								
543								
544								
545								

-In **Market Definition** table,the data of **Drug strength** is missing for some rows.

1	NDC	Drug Name	Brand/Generic Fla	Drug Form	Drug Strength	Drug Class
585	16110058780	CORDRAN TAPE	Brand	TAPE	N/A	Topical Corticosteroids
688	24338042001	PEDIADERM HC	Brand	KIT	N/A	Topical Corticosteroids
689	24338052001	PEDIADERM TA	Brand	KIT	N/A	Topical Corticosteroids
868	43538095310	TOVET	Brand	KIT	N/A	Topical Corticosteroids
1271	52544004424	CORDRAN TAPE	Brand	TAPE	N/A	Topical Corticosteroids
1272	52544004480	CORDRAN TAPE	Brand	TAPE	N/A	Topical Corticosteroids
1435	54288126001	ADV ALLERG COLLECT	Brand	KIT	N/A	Topical Corticosteroids
1602						
1603						
1604						
1605						
1606						
1607						
1608						

- Data of **Payers name** is missing in some of the entries in **Plan description**.

1	Plan ID	Plan Name	Payer name	Payer Channel
62	6811126	AFFORDABLE PPO PLATINUM SHP (MA)	-	-
243	22768999	MASS INSTITUTE OF TECHNOLOGY(MA)	-	COMMERCIAL
264	6793043	MERCYCARE EPO CO GOLD (WI)	-	-
283	9093247	MERCYCARE HIX EPO GNRL	-	COMMERCIAL
372	21481642	FED HIGH RISK POOL (SD)	-	-
229	6792956	FIRST HLTH PRMR PLS (MA)	-	MEDICAID
742	6812980	CIGNA-HLTHSPRING SECURE (ID)	-	MEDICAID
533	10079702	MUTUAL OF OMAHA RX PREMIER (WY)	-	MEDICARE
546				
547				
548				
549				

-Data of **Payer Channel** is missing for some of the rows in **Plan Description**.

A	B	C	D	E	F
Plan ID	Plan Name	Payer name	Payer Channel		
10025288	DEAN BRONZE PLUS (WI)	DEAN HEALTH PLAN (WI)	-		
6813306	BEST BUY COPAY HSA HMO BRNZ (MA)	HARVARD PILGRIM HEALTH CARE	-		
6807339	BEST BUY PPO GOLD SHOP (MA)	HARVARD PILGRIM HEALTH CARE	-		
6811126	AFFORDABLE PPO PLATINUM SHP (MA)	-	-		
6812567	ANTHEM HEALTH PLANS GENERAL	BCBS ANTHEM/WELLPOINT/WELLCHOICE	-		
6793043	MERCYCARE EPO CO GOLD (WI)	-	-		
6811347	MERCYCARE HMO SILVER SHOP (WI)	MERCYCARE HEALTH PLAN (WI)	-		
16967751	MERCYCARE HIX HMO GNRL	MERCYCARE HEALTH PLAN (WI)	-		
8276700	PRIORITY HEALTH (MI)	PRIORITY HEALTH (MI)	-		
21481642	FED HIGH RISK POOL (SD)	-	-		
6813552	CCHP PLATINUM SHOP (CA)	CHINESE COMMUNITY HLTH/CCHP (CA)	-		
6802985	DAKOTA CHOICE GOLD SHOP (SD)	DAKOTACARE (SD)	-		
16243639	LIFEGUARD (CA)	LIFEGUARD (CA)	-		
8005833	AETNA MEDICARE RX VALUE PLUS(TN)	AETNA INC	-		
10138548	CIGNA MEDICARE RX PLAN 3 (GA)	CIGNA	-		
6811545	CONNECTICARE VIP CUSTOM 1 (CT)	HIP/CONNECTICARE	-		
19371268	HEALTH NET PEARL OPTION 12 (CA)	HEALTH NET INC	-		
6793263	HEALTH NET PEARL OPT 9 (VA)	HEALTH NET INC	-		
6446167	ABRAZO ADVANTAGE GOLD (AZ)	PHOENIX MEMORIAL HOSPITAL (AZ)	-		

2) What % of patients move from diagnosis to treatment? What could be the probable reasons for some patients not getting on treatment?

Solution:

No of people diagnosed: **7261** (There are few people who are diagnosed 2 times)

No of people started treatment: **4507**

% of people moving from diagnosis to treatment: $4507 \times 100 / 7261 = \mathbf{62.071\%}$.

Probable reasons for some patients not getting on treatment :

- Financial barriers, such as lack of insurance or inability to afford the treatment.
- Lack of transportation to get to medical appointments.
- Limited access to healthcare providers in their area.
- Difficulty understanding or trusting the medical system.
- Fear or mistrust of the medical profession.
- Difficulty understanding the treatment or its potential side effects.
- Belief that the treatment is unnecessary or unproven.
- Co-occurring mental health or substance abuse issues.
- Cultural or linguistic barriers.
- Difficulty managing the treatment regimen.
- Personal beliefs or values that conflict with the treatment.

3) On average, how long (in days) do patients take to move from diagnosis to treatment?

Solution:

Undo		Clipboard		Font		Alignment		
G10								
	A	B	C	D	E	F	G	H
1	PatientID	Date of Diagnosis	Date of Treatment Started	Difference(Days)	Average			
2	100107674	11-08-2018	31-10-2018	81	274			
3	100352916	28-02-2018	14-05-2018	75				
4	100357133	08-03-2018	25-04-2018	48				
5	100378740	02-03-2018	14-03-2018	12				
6	100423759	19-03-2019	29-04-2019	41				
7	100517456	02-11-2017	01-11-2018	364				
8	100639327	20-12-2018	22-02-2019	64				
9	100640793	31-01-2017	10-01-2018	344				
10	101036153	08-03-2018	22-09-2018	198				
11	101092106	04-02-2018	28-06-2018	144				
12	101363189	15-05-2017	31-05-2017	16				
13	101363202	19-03-2017	06-12-2017	262				
14	101578090	10-04-2018	15-06-2018	66				
15	101690733	14-01-2017	14-01-2018	365				
16	101789657	30-03-2018	27-12-2018	272				
17	102091370	04-08-2017	07-08-2017	3				
18	102337591	27-06-2018	02-03-2019	248				
19	102379488	21-02-2018	03-04-2018	41				
20	102686832	29-03-2017	17-04-2018	384				
21	102914824	16-02-2017	16-02-2018	365				
22	103014611	16-01-2018	31-01-2018	15				
23	104011131	04-05-2017	26-06-2017	53				
24	104876896	17-09-2019	03-12-2019	77				
25	104963691	05-03-2017	30-03-2018	390				
26	105126596	18-07-2019	29-07-2019	11				
27	105154384	07-12-2018	04-02-2019	59				
28	105619000	06-08-2018	09-08-2018	3				

On average patient takes 274 days to move from Diagnosis to treatment.

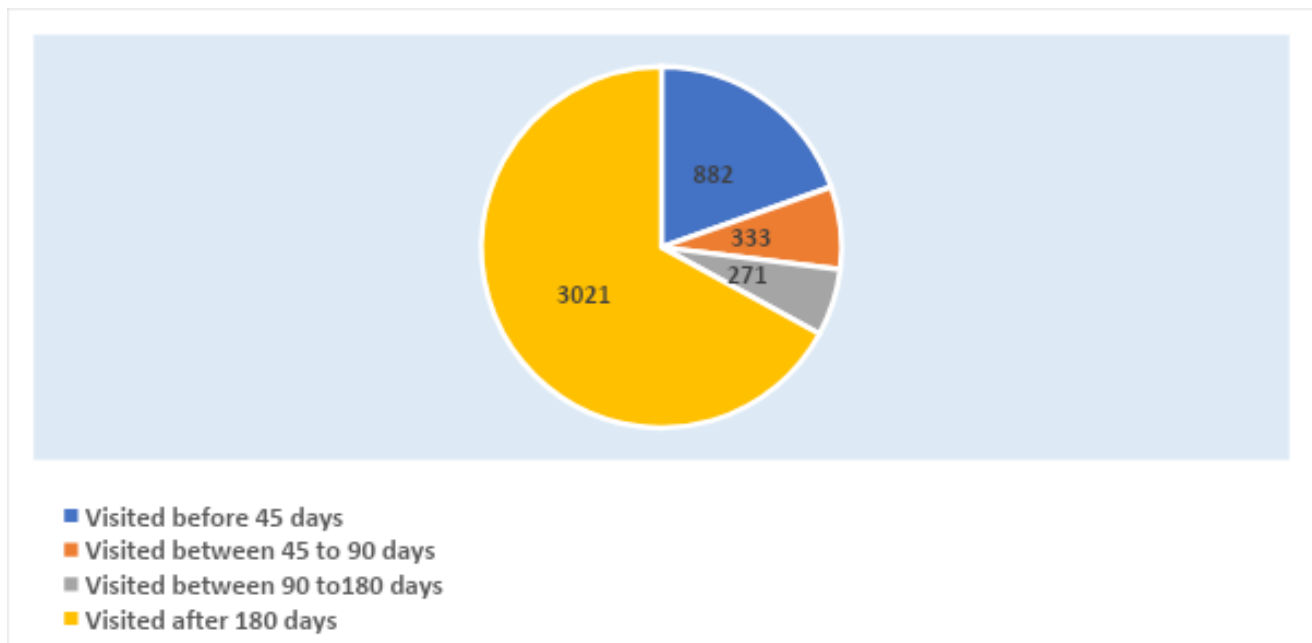
4) Create a graph to show patients started with the treatment within 45, 90, 180 and post 180 days.

Solution:

Table showing the required data

Days	No. of Peoples
Visited before 45 days	882
Visited between 45 to 90 days	333
Visited between 90 to 180 days	271
Visited after 180 days	3021
Grand Total	4507

-Corresponding Graph



5) What is the claim share of the payer channel over the years? Provide insights and visualizations. Do you see any change in claim pattern across years?

Solution:

-Claim share of Different payer channel over the years :

NO. of CLAIMS				
Payer Channel	2017	2018	2019	TOTAL
COMMERCIAL	6061	4626	3390	14077
MEDICAID	1154	962	1013	3129
MEDICARE	473	403	403	1279

-Insights and Visualization:

Top 10 Payers from Commercial Channel in 2018	
Payer Name	Count of Patient ID
AETNA INC	467
COVENTRY HEALTH CARE/FIRST HLTH	332
HEALTH NET INC	258
UNITED HEALTHCARE	248
BCBS ANTHEM/WELLPOINT/WELLCHOICE	186
HARVARD PILGRIM HEALTH CARE	154
MOLINA HEALTHCARE INC	150
MEDICAID	112
INDEPENDENT HEALTH ASSN (NY)	109
GEISINGER HEALTH PLAN	109

Top 10 Payers from Commercial Channel in 2019	
Payer Name	Count of Patient ID
AETNA INC	457
COVENTRY HEALTH CARE/FIRST HLTH	348
UNITED HEALTHCARE	162
HARVARD PILGRIM HEALTH CARE	150
MOLINA HEALTHCARE INC	147
HEALTH NET INC	137
BCBS ANTHEM/WELLPOINT/WELLCHOICE	116
CARESOURCE HEALTH PLAN	99
PROMINENCE HEALTH PLAN	86
COMMUNITYCARE (OK)	78

7) Considering Dermatology and Nurse Practitioners specialty physicians as the major targets for ProcDNA, which payer channel becomes more important and why?

Solution:

Commercial Channel becomes very important because the count of this is the largest for Dermatologist and Nurse practitioners

Payer Channel	Count of Payer.Payer Channel
COMMERCIAL	10036
MEDICAID	2219
MEDICARE	879
Grand Total	13134

