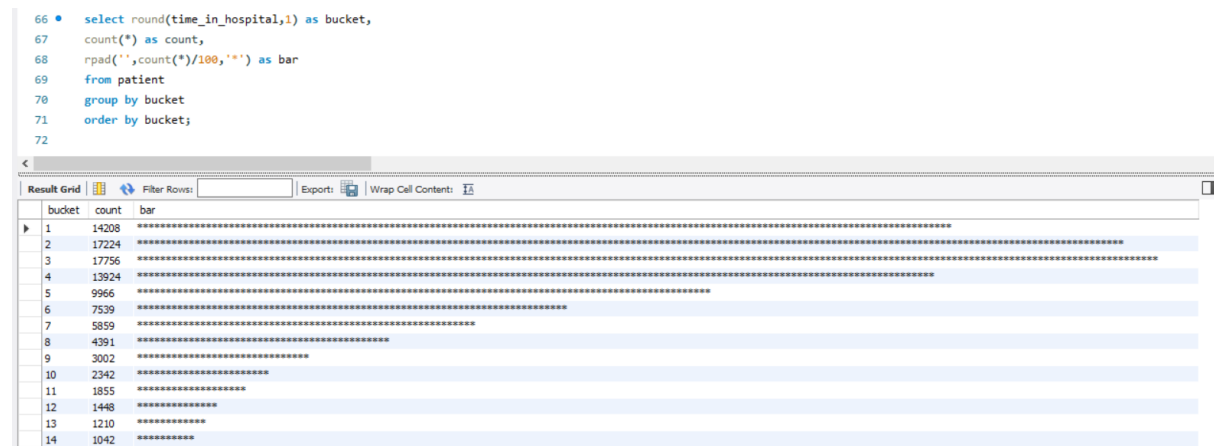


## HOSPITAL ANALYSIS

Even though SQL isn't primarily designed for creating visual charts, we were able to make a histogram using the "Time in Hospital" data. This helped us understand how long patients usually stay. After looking at the results, we found that the majority of patients stay for 7 days or less. Specifically, we noticed that the highest number of patients stayed for only 3 days before they were discharged.



For a business to do well, it's crucial to know what services it offers so that it can plan effectively. In our case, we wanted to figure out the different types of medical services the hospital provides. To do this, we sorted them based on their medical specialty. However, since the initial list only gave us a long list of specialties, we needed to use more commands to get a better and clearer understanding.

```

73 • select distinct(medical_specialty)
74     from patient
75     order by medical_specialty;
76

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	medical_specialty			
?				
	AllergyandImmunology			
	Anesthesiology			
	Anesthesiology-Pediatric			
	Cardiology			
	Cardiology-Pediatric			
	DCPTEAM			
	Dentistry			
	Dermatology			
	Emergency/Trauma			
	Endocrinology			
	Endocrinology-Metabolism			
	Family/GeneralPractice			
	Gastroenterology			
	Gynecology			
	Hematology			
	Hematology/Oncology			
	Hospitalist			
	InfectiousDiseases			
	InternalMedicine			
	Nephrology			
	Neurology			
	Neurophysiology			
	Obsterics&Gynecology-...			
	Obstetrics			
	ObstetricsandGynecology			

By using the "HAVING" and "ROUND" commands, we improved our understanding of those particular medical specialties. To dig even deeper, we specifically looked for medical specialties with an average greater than 2.5 and a count exceeding 50. Based on the results, we found that "Surgery-Thoracic" had the highest average number of procedures among all the medical specialties.

```

73 • select distinct(medical_specialty)
74   from patient
75   order by medical_specialty;
76
77 • select medical_specialty,round(AVG(num_procedures),1) as Avg_procedures,
78       count(*) as count
79   from patient group by medical_specialty
80   having count > 50 and avg_procedures > 2.5
81   order by avg_procedures desc;

```

medical_specialty	Avg_procedures	count
Surgery-Thoracic	3.5	109
Radiologist	3.2	1140
Surgery-Cardiovascular/Thoracic	3.2	652
Cardiology	2.7	5352
Surgery-Vascular	2.6	533

After analysing the patient data, our next goal was to inform our boss about the race with the highest average number of lab procedures. We used a command to calculate the average of "Number of Lab Procedures" for each race in the dataset and arranged them in descending order based on this average. The findings led us to the conclusion that African Americans have the highest average number of lab procedures among the different races.

```

84 • select race ,avg(num_lab_procedures) as avg_num_lab_procedures from patient
85   group by race order by avg_num_lab_procedures desc;

```

race	avg_num_lab_procedures
?	44.1047
AfricanAmerican	44.0851
Other	43.4369
Caucasian	42.8329
Hispanic	42.7933
Asian	41.2122

Next we took on the Command "UNION" this allows us to add rows together from two different data sets. The command below we

created a UNION where the race was African American and UP in the patient table. The results below are not fancy, but it outlines the Patient Numbers for African Americans.

```

84 • select race ,avg(num_lab_procedures) as avg_num_lab_procedures from patient
85   group by race order by avg_num_lab_procedures desc;
86
87 • select patient_nbr from patient where race ="AfricanAmerican"
88   union
89   select patient_nbr from patient where metformin ="up";
90
91

```

patient_nbr
86047875
89869032
77391171
49726791
86328819
92519352
107389323
63000108
62718876
56480238

Similar to our previous analysis with "HAVING," we examined the average number of procedures based on medical specialties. The difference this time is that we considered all results without filtering for counts greater than 50 and average procedures greater than 2.5. In these results, we found that Proctology had the highest average number of procedures, which was 4. However, it's essential to note that this high average is based on only one occurrence.

```

91 • select medical_specialty, round(avg(num_procedures),1) as avg_procedures,
92   count(*) as count
93   from patient group by medical_specialty order by avg_procedures desc;

```

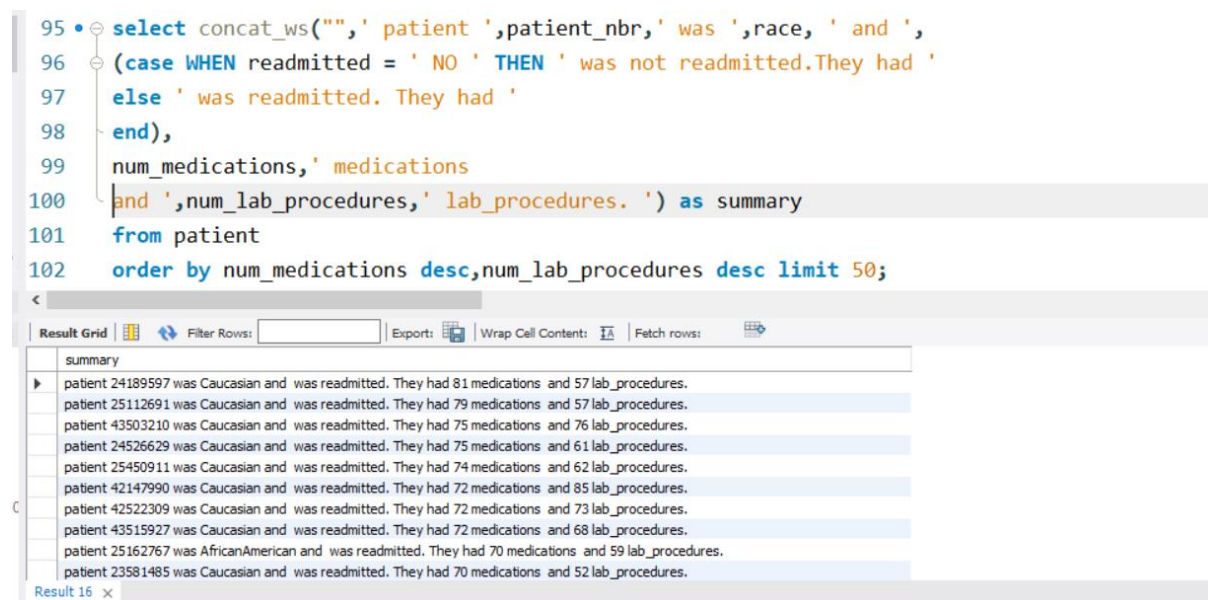
medical_specialty	avg_procedures	count
Proctology	4.0	1
Obstetrics&Gynecology-GynecologicOnco	3.6	25
Surgery-Thoracic	3.5	109
Surgery-Cardiovascular/Thoracic	3.2	652
Radiologist	3.2	1140
Dermatology	3.0	1
Surgery-PlasticwithInHeadandNeck	3.0	1
Cardiology	2.7	5352
Surgery-Vascular	2.6	533
Radiology	2.5	53

Our boss requested a list of patients who had been readmitted, along with the count of medications and lab procedures they had. To streamline this task, we utilized the "CONCAT" function, which enables us to combine text in our string to present the information as a sentence. In the command below, we used the Patient\_nbr from the patient dataset to retrieve our data. After that, we determined the number of medications and lab procedures so that we could include this information in our Concat statement. Additionally, we used INNER JOIN to ensure a smooth and connected process. We set a limit of 50 to manage the output efficiently.

```

95 • select concat_ws(""," patient ",patient_nbr,' was ',race, ' and ',
96 (case WHEN readmitted = ' NO ' THEN ' was not readmitted.They had '
97 else ' was readmitted. They had '
98 end),
99 num_medications,' medications
100 and ',num_lab_procedures,' lab_procedures. ') as summary
101 from patient
102 order by num_medications desc,num_lab_procedures desc limit 50;

```



The screenshot shows a SQL query editor with a query that uses the CONCAT function to create a summary string for each patient. The query selects patient information, readmission status, and counts of medications and lab procedures, ordered by these counts in descending order, limited to 50 results. Below the query, a 'Result Grid' displays the output as a table with one column named 'summary'. The table contains 16 rows of patient data, each formatted as a sentence describing the patient's race, readmission status, and the number of medications and lab procedures they had.

summary
patient 24189597 was Caucasian and was readmitted. They had 81 medications and 57 lab_procedures.
patient 25112691 was Caucasian and was readmitted. They had 79 medications and 57 lab_procedures.
patient 43503210 was Caucasian and was readmitted. They had 75 medications and 76 lab_procedures.
patient 24526629 was Caucasian and was readmitted. They had 75 medications and 61 lab_procedures.
patient 25450911 was Caucasian and was readmitted. They had 74 medications and 62 lab_procedures.
patient 42147990 was Caucasian and was readmitted. They had 72 medications and 85 lab_procedures.
patient 42522309 was Caucasian and was readmitted. They had 72 medications and 73 lab_procedures.
patient 43515927 was Caucasian and was readmitted. They had 72 medications and 68 lab_procedures.
patient 25162767 was AfricanAmerican and was readmitted. They had 70 medications and 59 lab_procedures.
patient 23581485 was Caucasian and was readmitted. They had 70 medications and 52 lab_procedures.

From the results above you can see each line has the Patient Number what their race is, and whether they were readmitted. Then it shows how many medications they have and the number of lab procedures performed for them. This will help our boss present the facts when determining how to best help the hospital.

We used a Subquery to help determine our results.

Results show the encounter ID along with time in hospital the medical specialty and the number of lab procedure for the patient.

```

104 • with avg_time as (select avg(time_in_hospital) from patient)
105   select * from patient
106   where admission_type_id = 1
107   and time_in_hospital > (select * from avg_time);

```

encounter_id	patient_nbr	race	gender	age	weight	admission_type_id	discharge_disposition_id	admission_source_id	time_in_hospital	payer_code	medical_specialty	num_lab_procedures	num_p
63768	114882984	Caucasian	Male	[70-80]	?	1	1	7	5	?	?	73	0
28236	89869032	AfricanAmerican	Female	[40-50]	?	1	1	7	9	?	?	47	2
40926	85504905	Caucasian	Female	[40-50]	?	1	3	7	7	?	Family/GeneralPractice	60	0
42570	77586282	Caucasian	Male	[80-90]	?	1	6	7	10	?	Family/GeneralPractice	55	1
73578	86328819	AfricanAmerican	Male	[60-70]	?	1	3	7	12	?	?	75	5
89682	107389323	AfricanAmerican	Male	[70-80]	?	1	1	7	5	?	?	35	5
236316	40523301	Caucasian	Male	[80-90]	?	1	3	7	6	?	Cardiology	64	3
252822	18196434	Caucasian	Female	[80-90]	?	1	2	7	5	?	Cardiology	52	0
253380	56480238	AfricanAmerican	Female	[60-70]	?	1	1	7	6	?	?	87	0
260166	80845353	Caucasian	Female	[70-80]	?	1	1	7	6	?	Family/GeneralPractice	27	0
358776	101002446	Caucasian	Male	[70-80]	?	1	6	7	7	?	?	47	2
450210	80177094	Caucasian	Female	[80-90]	?	1	11	7	7	?	?	72	1
590346	37746639	AfricanAmerican	Female	[40-50]	?	1	6	7	8	?	?	55	2
870294	95075649	Caucasian	Female	[70-80]	?	1	6	7	7	?	?	75	2
881016	55152216	Caucasian	Male	[50-60]	?	1	1	7	12	?	?	83	2
955884	93196251	Caucasian	Female	[70-80]	?	1	3	7	5	?	Cardiology	34	0
1079592	101707335	Caucasian	Female	[50-60]	?	1	7	7	5	?	?	58	0
1161024	20830941	Caucasian	Female	[70-80]	?	1	2	7	5	?	Cardiology	36	0
1194678	22997880	AfricanAmerican	Female	[60-70]	?	1	6	7	6	?	?	65	3
1195548	106343838	Caucasian	Female	[70-80]	?	1	3	7	5	?	Family/GeneralPractice	64	0
1260312	38688588	Caucasian	Male	[70-80]	?	1	11	7	9	?	?	56	2

## Summary:

The healthcare field is really important for society, and those in it play a crucial role in providing help when needed. This project focused on a small aspect of healthcare, and I discovered some interesting things. Most patients tend to stay in the hospital for less than 7 days, with the majority staying between two and three days. I also found out that the medical specialty with the highest average number of procedures is Surgery-Thoracic. Additionally, I learned that African Americans have the highest average number of lab procedures. What's cool is that through SQL, you can create commands that generate sentences using both data and text. This saved a lot of time because I didn't have to write out each patient's information individually, especially regarding whether they were readmitted or not.

