



# BANCO DE DADOS ADAPR3

# CONTENT

**01**

OUR TEAM

**02**

BUSINESS UNDERSTANDING

**03**

GOALS AND OBJETIVOS

**04**

BANCO DE DADOS

**05**

DATA INSIGHTS

**06**

MACHINE LEARNING MODEL

**07**

GREETINGS

# OUR TEAM



MAURÍCIO SOBRINHO - HEAD DE DADOS

ANE CAROLINE - 1115018 DATA SCIENTIST

CHEN YEN PIN - 1115013 DATA SCIENTIST

GABRIEL ALVARENGA 1115XXX DATA SCIENTIST

GABRIEL VILARIHO- 115016DATA SCIENTIST

GABRIELA MENDES - 1115XXX DATA SCIENTIST

JESYA LIMA- 1115011 DATA SCIENTIST

MARCEL CLERES - 1115005 DATA SCIENTIST

RENAN SILVA - 1115007 DATA SCIENTIST





# GOALS AND OBJECTIVES

## Objective n° 1

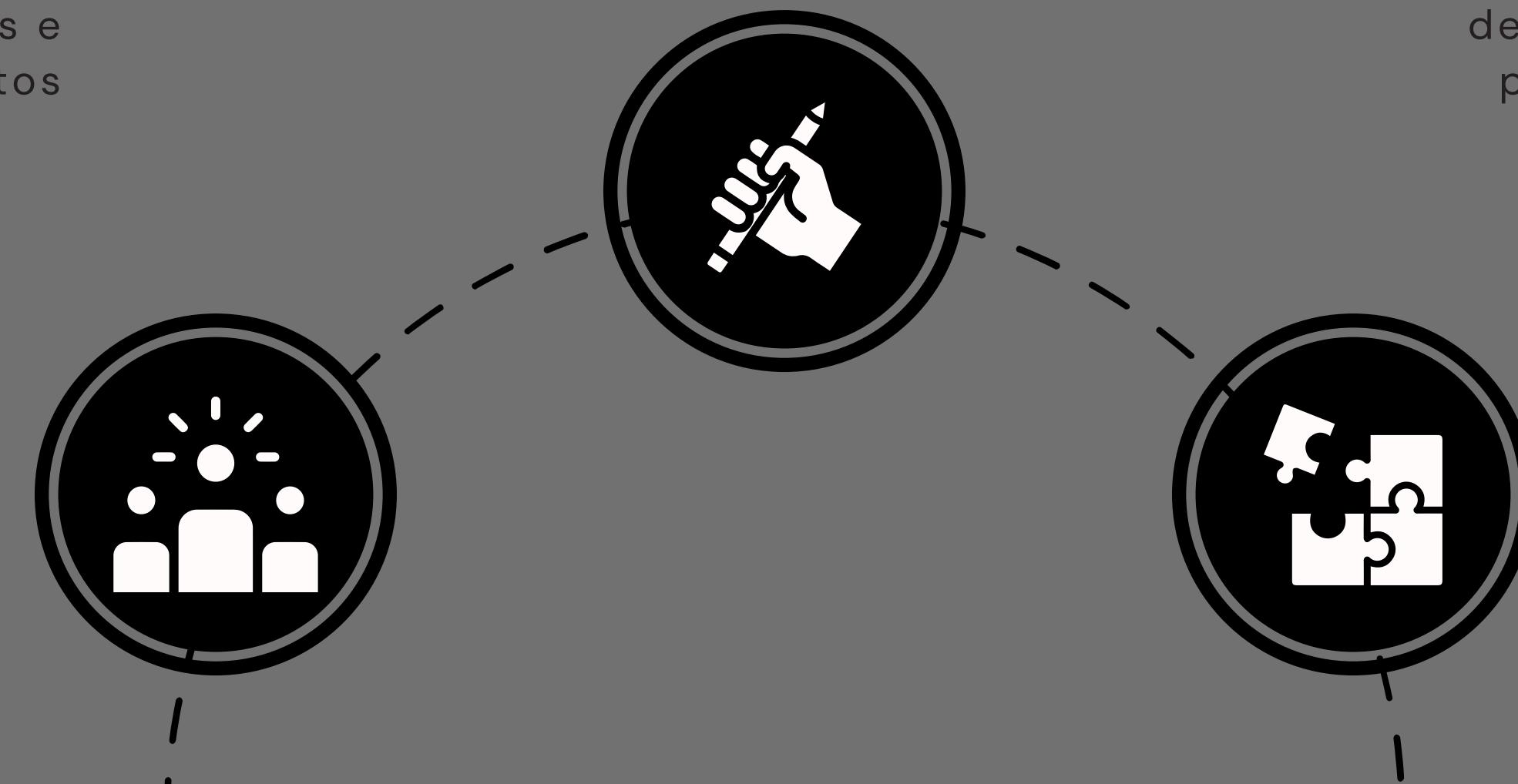
Criação de um banco de dados, entendendo as suas entidades , atributos cardinalidades e relacionamentos

## Objective n° 2

Usar linguagem SQL no SGBD POSTGRESQL.  
Gerar consultas(querys), extrair informações e insights.

## Objective n° 3

Usar modelo de Machine Learning para predizer possíveis níveis de satisfação, com isso prever futuras ações de retenção de clientes



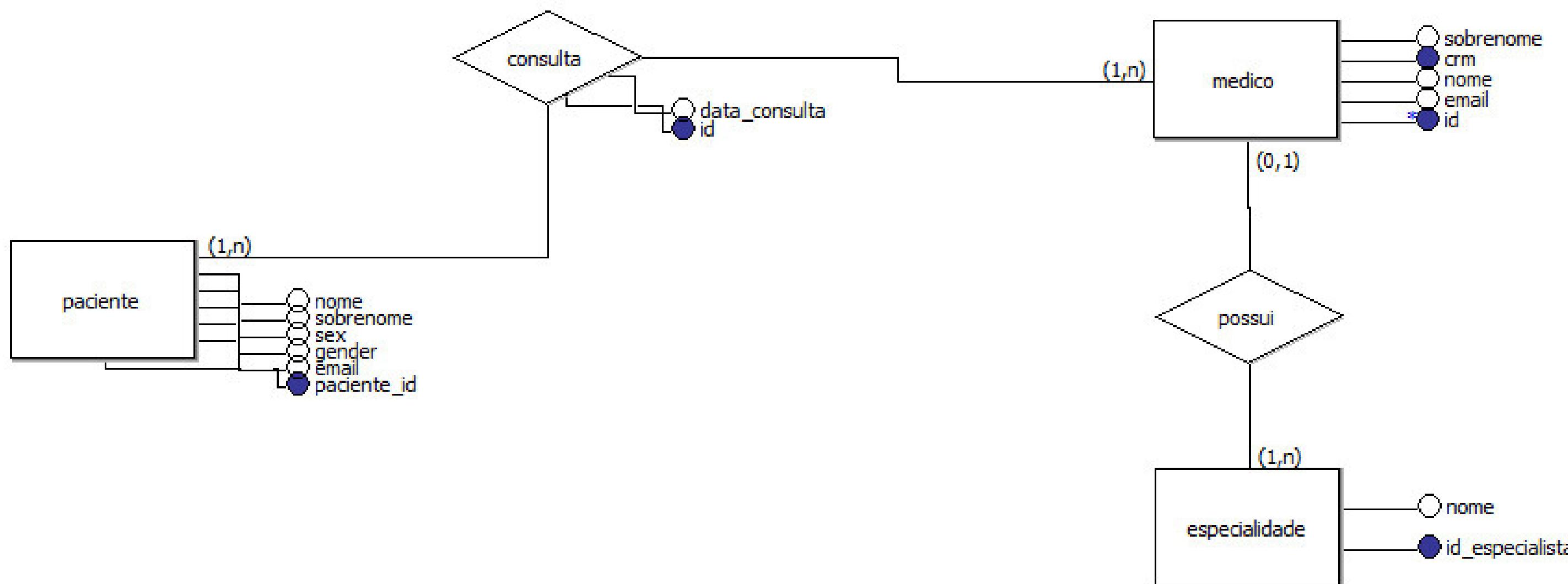
# BUSINESS UNDERSTANDING

- New Era Hospital incorporou um hospital na região de São Carlos, porém ao analisar o banco de dados existente, deparou-se com o armazenamento sendo feito em planilhas do Excel
- Também constatou que há uma parcela de clientes que sentem insatisfeitas com o atendimento prestado e usaram portais de reclamação para levantar o alcance da insatisfação
- Para organizar e gerar mais valor aos dados foi solicitado a Ada Data Consulting um projeto para criar e melhorar um banco de dados e também um modelo de classificação de churn



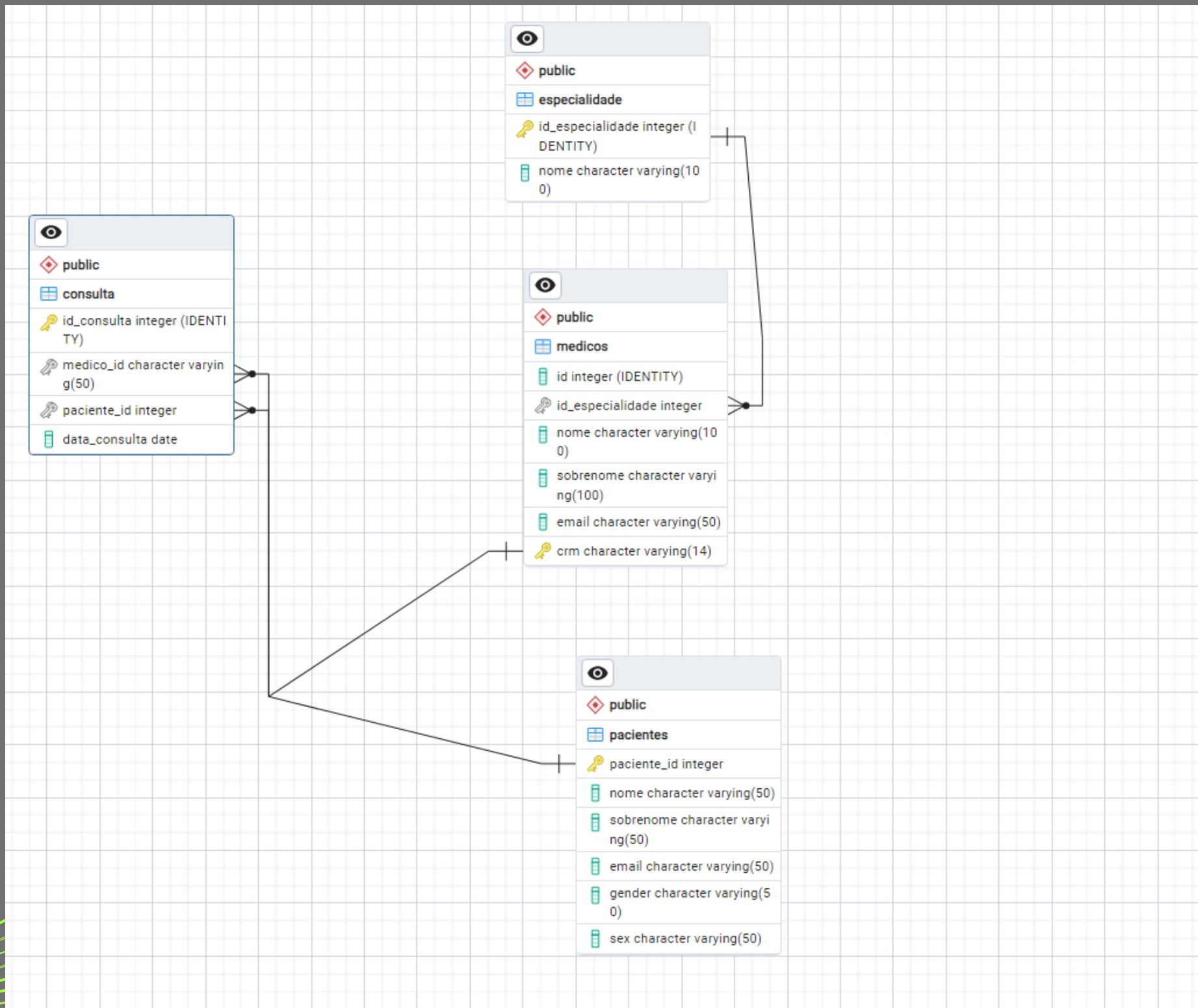
# BANCO DE DADOS

Desenho esquemáticos



# BANCO DE DADOS

Carinadalidades



# BANCO DE DADOS

Resultado final

```
select * from especialidade limit 5;
```

```
1 select * from pacientes limit 5;
```

Data Output Messages Notifications



id_especialidade [PK] integer	nome character varying
1	Ginecologista
2	Pediatra
3	Cirurgiao Plastico
4	Psicologo
5	Cardiologista

Data Output Messages Notifications



paciente_id [PK] integer	nome character varying	sobrenome character varying	email character varying	gender character varying	sex character varying	idade integer
1	Peter	Trude	ptrude0@merriam-webster.com	Male	Female	62
2	Adena	Meagher	ameagher1@mediafire.com	Female	Male	90
3	Quinlan	Spittle	qspittle2@technorati.com	Male	Female	82
4	Clayborn	Killwick	ckillwick3@noaa.gov	Male	Female	41
5	Jonie	Barnfield	jbarnfield4@youtu.be	Female	Male	60

# BANCO DE DADOS

Resultado final

```
select * from consulta limit 5;
```

Output    Messages    Notifications

File    Save    Copy    Paste    Delete    Database    Download    Refresh

<b>id_consulta</b> [PK] integer	<b>medico_id</b> character varying	<b>paciente_id</b> integer	<b>data_consulta</b> date
1	100035/SP	203	2024-02-26
2	100049/RJ	385	2023-03-30
3	100015/SP	211	2023-03-16
4	100070/SP	994	2023-05-02
5	100074/SP	99	2024-02-22

```
select * from medico limit 5;
```

Output    Messages    Notifications

File    Save    Copy    Paste    Delete    Database    Download    Refresh

<b>id</b> integer	<b>id_especialidade</b> integer	<b>nome</b> character varying	<b>sobrenome</b> character varying	<b>email</b> character varying	<b>crm</b> [PK] character varying
1	9	Carmelia	Falck	cfalck0@xinhuanet.com	100001/SP
2	12	Valencia	De Witt	vdewitt1@amazon.com	100010/RJ
3	5	Shara	Chetwind	schetwind2@freewebs.com	100024/SP
4	6	Pattie	Dulanty	pdulanty3@dell.com	100093/SP
5	11	Micky	Dencs	mdencs4@latimes.com	100002/RJ

# DATA INSIGHTS

## Querys e consultas

```
43 -- Vendo a quantidade de atendimento por gênero
44 select
45     pc.gender,
46     count(cs.id_consulta) as qtd_consultas,
47     ROUND(COUNT(cs.id_consulta) * 100.0 / SUM(COUNT(cs.id_consulta)) OVER(), 2) AS porcentagem
48 from
49     pacientes as pc
50 inner join
51     consulta as cs on cs.paciente_id = pc.paciente_id
52 group by
53     pc.gender
54 order by
55     qtd_consultas desc;
56
```

Data Output Messages Notifications



	gender character varying	qtd_consultas bigint	porcentagem numeric
1	Male	4650	46.50
2	Female	4344	43.44
3	Bigender	212	2.12
4	Genderfluid	186	1.86
5	Agender	166	1.66
6	Genderqueer	158	1.58
7	Non-binary	144	1.44
8	Polygender	140	1.40

# DATA INSIGHTS

Querys e consultas

```
58 -- Vendo a quantidade de atendimento por sexo
59 select
60     pc.sex,
61     count(cs.id_consulta) as qtd_consultas,
62     ROUND(COUNT(cs.id_consulta) * 100.0 / SUM(COUNT(cs.id_consulta))) OVER(), 2) AS porcentagem
63 from
64     pacientes as pc
65 inner join
66     consulta as cs on cs.paciente_id = pc.paciente_id
67 group by
68     pc.sex;
69
70
```

Data Output Messages Notifications



	sex character varying	qtd_consultas bigint	porcentagem numeric
1	Female	5077	50.77
2	Male	4923	49.23

# DATA INSIGHTS

## Querys e consultas

```
71 -- Verificar a quantidade de consultas por especialidade
72 SELECT
73     esp.nome AS especialidade,
74     COUNT(cs.id_consulta) AS qtd_consultas,
75     ROUND(COUNT(cs.id_consulta) * 100.0 / SUM(COUNT(cs.id_consulta)) OVER(), 2) AS porcentagem
76 FROM
77     especialidade AS esp
78 INNER JOIN
79     medico AS md ON esp.id_especialidade = md.id_especialidade
80 INNER JOIN
81     consulta AS cs ON md.crm = cs.medico_id
82 GROUP BY
83     esp.nome
84 ORDER BY
85     qtd_consultas DESC;
```

Data Output    Messages    Notifications

	especialidade character varying	qtd_consultas bigint	porcentagem numeric
1	Cirurgiao Plastico	2414	24.14
2	Cardiologista	1428	14.28
3	Pediatra	1374	13.74
4	Oftalmologia	1373	13.73
5	Psicologo	1044	10.44
6	Neurologia	636	6.36
7	Oncologia	365	3.65
8	Ginecologista	358	3.58

# DATA INSIGHTS

## Querys e consultas

```
88 -- Fazendo a classificações de consultas já realizadas ou futuras
89 select
90     cs.id_consulta,
91     cs.medico_id,
92     cs.paciente_id,
93     cs.data_consulta,
94     case
95         when cs.data_consulta < current_date then 'Consulta Realizada'
96         else 'Consulta Futura'
97     end as status_consulta
98 from
99     consulta as cs;
100
101
```

Data Output Messages Notifications

The screenshot shows a database query results interface. At the top, there is a code editor window displaying the SQL query. Below it is a toolbar with various icons for file operations. The main area is a table with the following data:

	id_consulta [PK] integer	medico_id character varying	paciente_id integer	data_consulta date	status_consulta text
1	2	100035/SP	203	2024-02-26	Consulta Realizada
2	3	100049/RJ	385	2023-03-30	Consulta Realizada
3	4	100015/SP	211	2023-03-16	Consulta Realizada
4	5	100070/SP	994	2023-05-02	Consulta Realizada
5	6	100074/SP	99	2024-02-22	Consulta Realizada
6	7	100035/SP	238	2024-01-01	Consulta Realizada
7	8	100092/SP	253	2024-01-04	Consulta Realizada
8	9	100033/RJ	742	2023-07-31	Consulta Realizada
9	10	100002/SP	592	2024-02-02	Consulta Realizada

Total rows: 1000 of 100000 Query complete: 00:00:00.448

# DATA INSIGHTS

## Querys e consultas

```
102 -- Relação de consulta x sexo_paciente x especialidade_medico
103 select
104     pc.sex as sexo_paciente,
105     esp.nome as especialidade_do_medico,
106     count(cs.id_consulta) as total_atendimentos,
107     ROUND(COUNT(cs.id_consulta) * 100.0 / SUM(COUNT(cs.id_consulta)) OVER(), 2) AS porcentagem
108 from
109     consulta as cs
110 inner join
111     pacientes AS pc ON cs.paciente_id = pc.paciente_id
112 INNER JOIN
113     medico AS md ON cs.medico_id = md.crm
114 INNER JOIN
115     especialidade AS esp ON md.id_especialidade = esp.id_especialidade
116 GROUP BY
117     esp.nome,
118     pc.sex
119 ORDER BY
120     especialidade_do_medico;
121
```

Data Output Messages Notifications



	sexo_paciente character varying	especialidade_do_medico character varying	total_atendimentos bigint	porcentagem numeric
1	Male	Cardiologista	681	6.81
2	Female	Cardiologista	747	7.47
3	Female	Cirurgiao Plastico	1201	12.01
4	Male	Cirurgiao Plastico	1213	12.13
5	Male	Endocrinologia	140	1.40
-	-	-	-	-

# DATA INSIGHTS

## Querys e consultas

```
123 -- Classificando o público pela idade
124
125 select
126   case
127     when idade < 30 then 'Jovem'
128     when idade between 30 and 60 then 'Adulto'
129     else 'Idoso'
130   end as faixa_etaria_paciente,
131   count(id_consulta) as total_consultas,
132   ROUND(COUNT(cs.id_consulta) * 100.0 / SUM(COUNT(cs.id_consulta)) OVER(), 2) AS porcentagem
133 from
134   pacientes as pc
135   INNER JOIN
136     consulta as cs on pc.paciente_id = cs.paciente_id
137   group by
138     faixa_etaria_paciente;
```

Data Output    Messages    Notifications



	faixa_etaria_paciente	total_consultas	porcentagem
	text	bigint	numeric
1	Jovem	1107	11.07
2	Idoso	5346	53.46
3	Adulto	3547	35.47

# DATA INSIGHTS

## Querys e consultas

```
141 -- Vendo a dispersão do público por área de atendimento médica
142
143 select
144     esp.nome as especialidade_do_medico,
145     case
146         when idade < 30 then 'Jovem'
147         when idade between 30 and 60 then 'Adulto'
148         else 'Idoso'
149     end as faixa_etaria_paciente,
150     count(id_consulta) as total_consultas,
151     ROUND(count(cs.id_consulta) * 100.0 / sum(count(cs.id_consulta)) over(partition by esp.nome), 2) as porcentagem
152 from
153     pacientes as pc
154 inner join
155     consulta as cs on pc.paciente_id = cs.paciente_id
156 inner join
157     medico as md on cs.medico_id = md.crm
158 inner join
159     especialidade as esp on md.id_especialidade = esp.id_especialidade
160 group by
161     especialidade_do_medico,
162     faixa_etaria_paciente
163 order by
164     especialidade_do_medico;
165
```

Data Output Messages Notifications



	especialidade_do_medico	faixa_etaria_paciente	total_consultas	porcentagem
1	Cardiologista	Adulto	486	34.03
2	Cardiologista	Idoso	775	54.27
3	Cardiologista	Jovem	167	11.69
4	Cirurgiao Plastico	Adulto	856	35.46
5	Cirurgiao Plastico	Idoso	1303	53.98
6	Cirurgiao Plastico	Jovem	255	10.56
7	Endocrinologia	Adulto	108	33.64
8	Endocrinologia	Idoso	176	54.83
9	Endocrinologia	Jovem	37	11.53
10	Ginecologista	Adulto	131	36.59
11	Ginecologista	Idoso	193	53.91
12	Ginecologista	Jovem	24	0.50

Total rows: 22 of 22 Query complete 00:00:00.001

# DATA INSIGHTS

## Querys e consultas

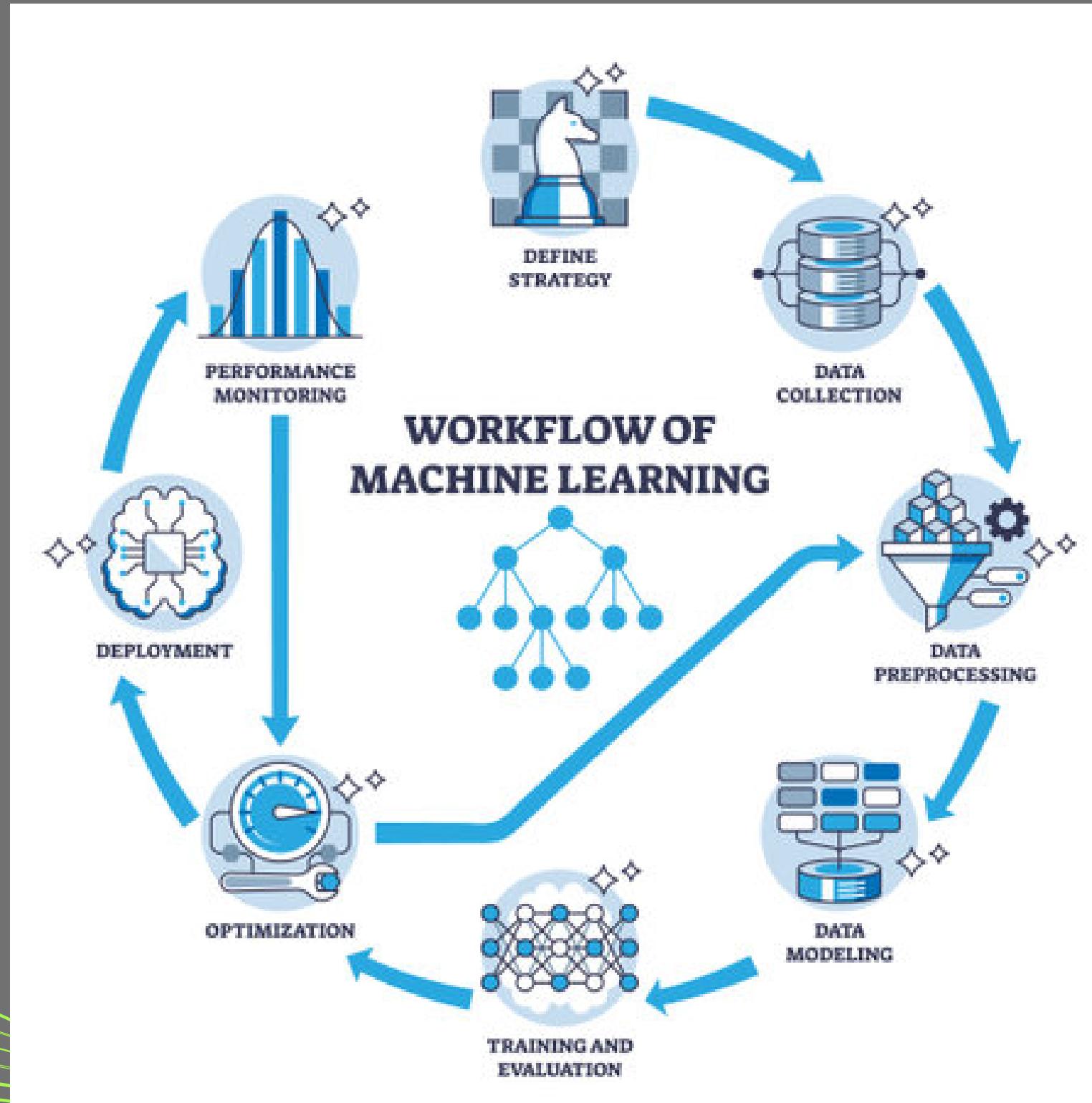
```
167 -- Ver o total de consultas por médicos
168 select
169     md.nome,
170     md.crm,
171     count(cs.id_consulta) as total_consultas
172 from
173     medico as md
174 inner join
175     consulta as cs on md.crm = cs.medico_id
176 group by
177     md.nome,
178     md.crm
179 order by
180     total_consultas desc;
181
```

Data Output Messages Notifications

	nome character varying	crm [PK] character varying	total_consultas bigint
1	Gerome	100085/SP	687
2	Shara	100024/SP	676
3	Brynn	100038/RJ	658
4	Jenine	100079/SP	394
5	Peta	100017/SP	385
6	Tori	100090/SP	366
7	Erwin	100033/RJ	365
8	Husein	100031/RJ	364
9	Matthus	100038/SP	364
10	Vanny	100074/SP	358
11	Zelda	100092/SP	358

# MACHINE LEARNING MODEL

## Workflow



# MACHINE LEARNING MODEL

Etapas para a criação do modelo até o deploy

- Criando ambiente virtual conda create e determinando os requirement necessários
- Importando dados
- Preprocessamento e treinamento do modelo usando PIPELINE E COLUMN TRANSFORM para garantir integridade do modelo
- Avaliando os modelos
- Escolhendo os modelos com melhores resultados
- Criar modelo MVP no streamlit
- Deploy local e aprovando subir para uso da equipe



**GIT CLONE**  
**HTTPS://GITHUB.COM/CHENYP-**  
**1982/ADAPR3**

# DEPLOY LOCALLY



localhost:8501

medico\_id  
100035/SP

O medico escolhido foi 100035/SP

sexo\_paciente  
Male

Voce selecionou: Male

especialidade  
Ginecologista

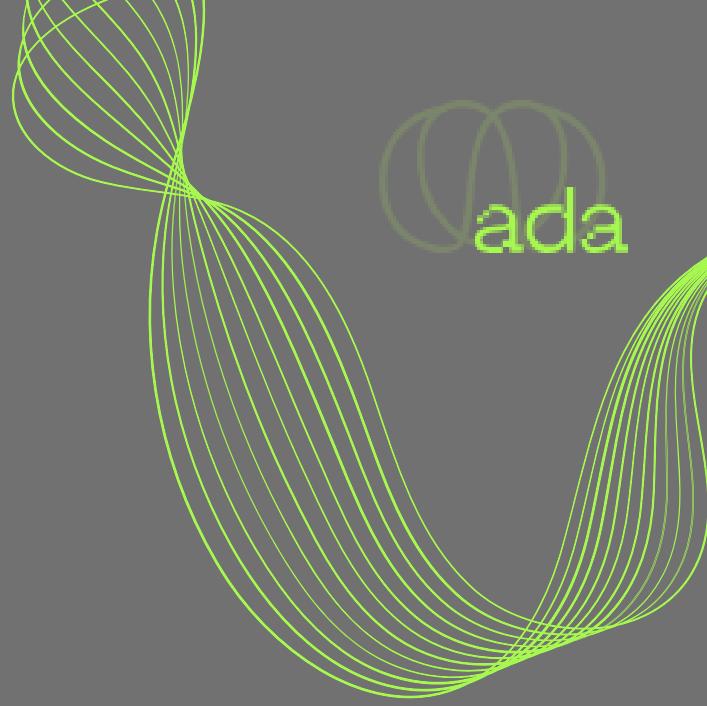
idade  
30

Voce selecionou: 30

Predict

A possivel recomendacao e 1

New Chrome available



# TESTIMONIALS

**Rufus Stewart**



Depois que foi implementado o sistema de banco de dados, ficou mais facil de encontrarmos as informações dos pacientes



**Morgan Maxwell**



O modelo de machine learning veio em uma otima hora para conseguirmos abordar os pacientes e melhorar o atendimento!!!



# MUITO OBRIGADO

*“ Não importa quanto você aprendeu,  
mas quanto absorveu daquilo que  
aprendeu.”*

Bruce Lee



ada