

## *Dental carries detection from dental x-rays*

CSV.FILE

Count y	Calen dar Year	Measure	Age Filt er	User s	Denomin ator (3 Months Continuo us Eligibility)	Utilizat ion %
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age <1	164	6,878	2.38%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 10- 14	16,6 38	33,609	49.50%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 1-2	4,66 5	16,442	28.37%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 15- 18	10,1 16	24,864	40.69%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 19- 20	2,67 8	9,457	28.32%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 21- 34	3,55 9	31,209	11.40%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 3-5	13,6 55	25,134	54.33%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 35- 44	1,76 7	19,426	9.10%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 45- 64	3,92 4	36,231	10.83%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 65- 74	1,79 4	21,691	8.27%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 6-9	17,8 47	31,674	56.35%
Alame da	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 75+	1,87 3	22,002	8.51%
Alame da	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age <1	0	6,878	0.00%
Alame da	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 10- 14	11,7 00	33,609	34.81%
Alame da	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 1-2		16,442	

Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 15-18	7,646	24,864	30.75%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 19-20	2,046	9,457	21.63%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 21-34	737	31,209	2.36%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 3-5	8,086	25,134	32.17%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 35-44		19,426	
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 45-64	424	36,231	1.17%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 65-74	200	21,691	0.92%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 6-9	11,796	31,674	37.24%
Alameda	CY 2013	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 75+	358	22,002	1.63%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age <1		6,878	
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 10-14	5,614	33,609	16.70%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 1-2		16,442	
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 15-18	3,908	24,864	15.72%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 19-20	1,228	9,457	12.99%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 21-34	1,889	31,209	6.05%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 3-5	3,275	25,134	13.03%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 35-44	1,011	19,426	5.20%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 45-64	2,373	36,231	6.55%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 65-74	1,091	21,691	5.03%

Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 6-9	6,550	31,674	20.68%
Alameda	CY 2013	Use of Dental Treatment Services (D2000 - D9999)	Age 75+	1,420	22,002	6.45%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age <1	85	6,878	1.24%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 10-14	11,687	33,609	34.77%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 1-2	1,491	16,442	9.07%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 15-18	7,429	24,864	29.88%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 19-20	1,925	9,457	20.36%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 21-34	756	31,209	2.42%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 3-5	7,471	25,134	29.72%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 35-44	139	19,426	0.72%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 45-64	282	36,231	0.78%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 65-74	111	21,691	0.51%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 6-9	11,733	31,674	37.04%
Alameda	CY 2013	Use of Preventive Services (D1000 - D1999)	Age 75+	204	22,002	0.93%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age <1	0	6,878	0.00%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 10-14	4,029	33,609	11.99%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 1-2	187	16,442	1.14%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 15-18	3,027	24,864	12.17%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 19-20	869	9,457	9.19%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 21-34	372	31,209	1.19%

Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 3-5	2,816	25,134	11.20%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 35-44	65	19,426	0.33%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 45-64	108	36,231	0.30%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 65-74	18	21,691	0.08%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 6-9	5,053	31,674	15.95%
Alameda	CY 2013	Use of Restorative Services (D2000 - D2999)	Age 75+	27	22,002	0.12%
Alameda	CY 2013	Use of Sealant (D1351)	Age <1		6,878	
Alameda	CY 2013	Use of Sealant (D1351)	Age 10-14	2,113	33,609	6.29%
Alameda	CY 2013	Use of Sealant (D1351)	Age 1-2		16,442	
Alameda	CY 2013	Use of Sealant (D1351)	Age 15-18		24,864	
Alameda	CY 2013	Use of Sealant (D1351)	Age 19-20		9,457	
Alameda	CY 2013	Use of Sealant (D1351)	Age 21-34		31,209	
Alameda	CY 2013	Use of Sealant (D1351)	Age 3-5		25,134	
Alameda	CY 2013	Use of Sealant (D1351)	Age 35-44		19,426	
Alameda	CY 2013	Use of Sealant (D1351)	Age 45-64		36,231	
Alameda	CY 2013	Use of Sealant (D1351)	Age 65-74		21,691	
Alameda	CY 2013	Use of Sealant (D1351)	Age 6-9	3,831	31,674	12.10%
Alameda	CY 2013	Use of Sealant (D1351)	Age 75+		22,002	
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age <1	81	6,878	1.18%

Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 10-14	11,055	33,609	32.89%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 1-2	1,404	16,442	8.54%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 15-18	6,705	24,864	26.97%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 19-20	1,693	9,457	17.90%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 21-34	580	31,209	1.86%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 3-5	7,453	25,134	29.65%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 35-44	94	19,426	0.48%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 45-64	230	36,231	0.63%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 65-74	102	21,691	0.47%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 6-9	11,637	31,674	36.74%
Alameda	CY 2013	Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351)	Age 75+	184	22,002	0.84%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age <1	97	7,325	1.32%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 10-14	18,331	36,732	49.90%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 1-2	4,742	16,868	28.11%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 15-18	10,659	27,061	39.39%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 19-20	2,882	11,642	24.76%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 21-34	9,969	69,229	14.40%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 3-5	13,659	25,180	54.25%

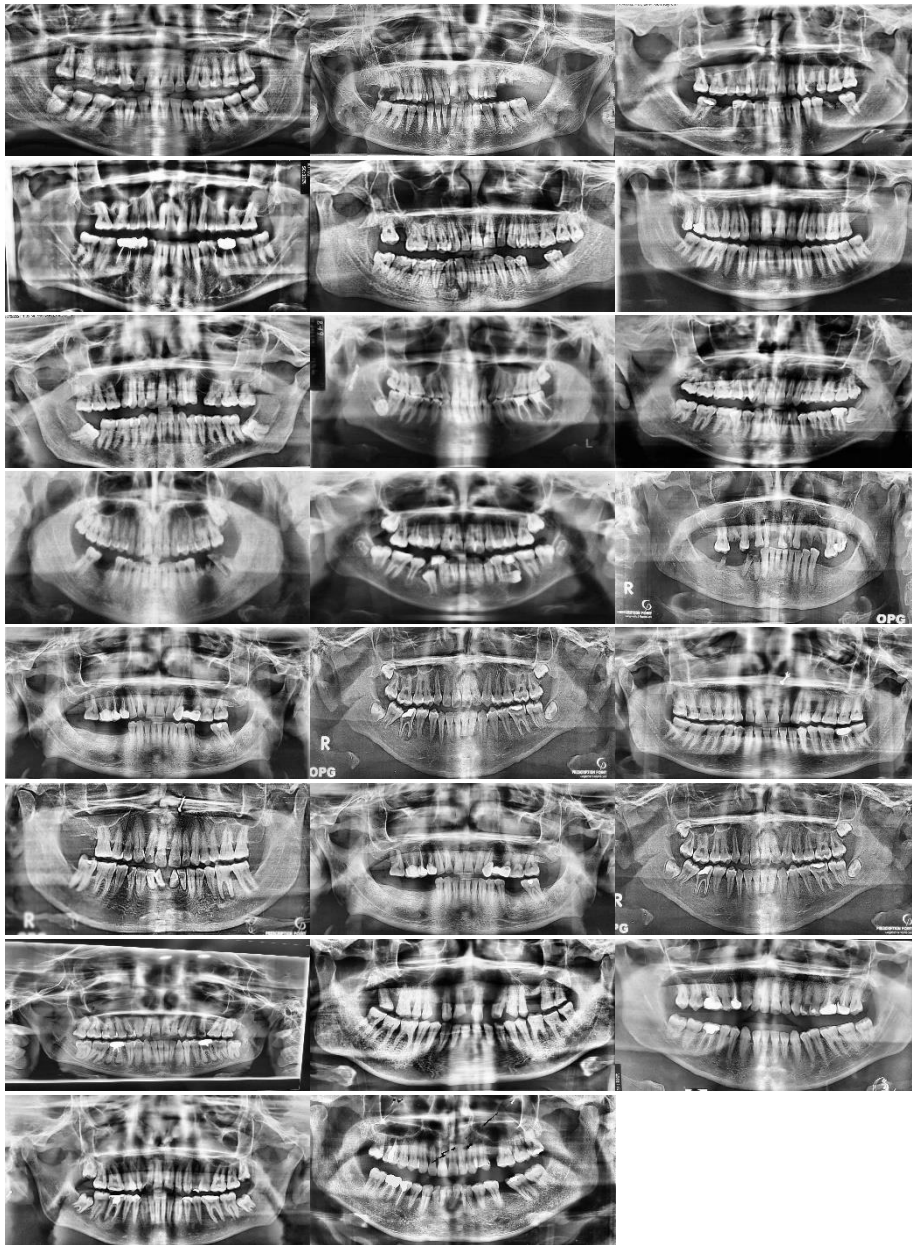
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 35-44	6,075	34,481	17.62%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 45-64	16,363	78,303	20.90%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 65-74	5,183	23,543	22.02%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 6-9	19,476	34,024	57.24%
Alameda	CY 2014	Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)	Age 75+	4,281	22,559	18.98%
Alameda	CY 2014	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age <1	0	7,325	0.00%
Alameda	CY 2014	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 10-14	12,379	36,732	33.70%
Alameda	CY 2014	Exams/Oral Health Evaluations (D0120, D0145, or D0150)	Age 1-2	13	16,868	0.08%

IMAGE DATASET:

Impacted teeth:

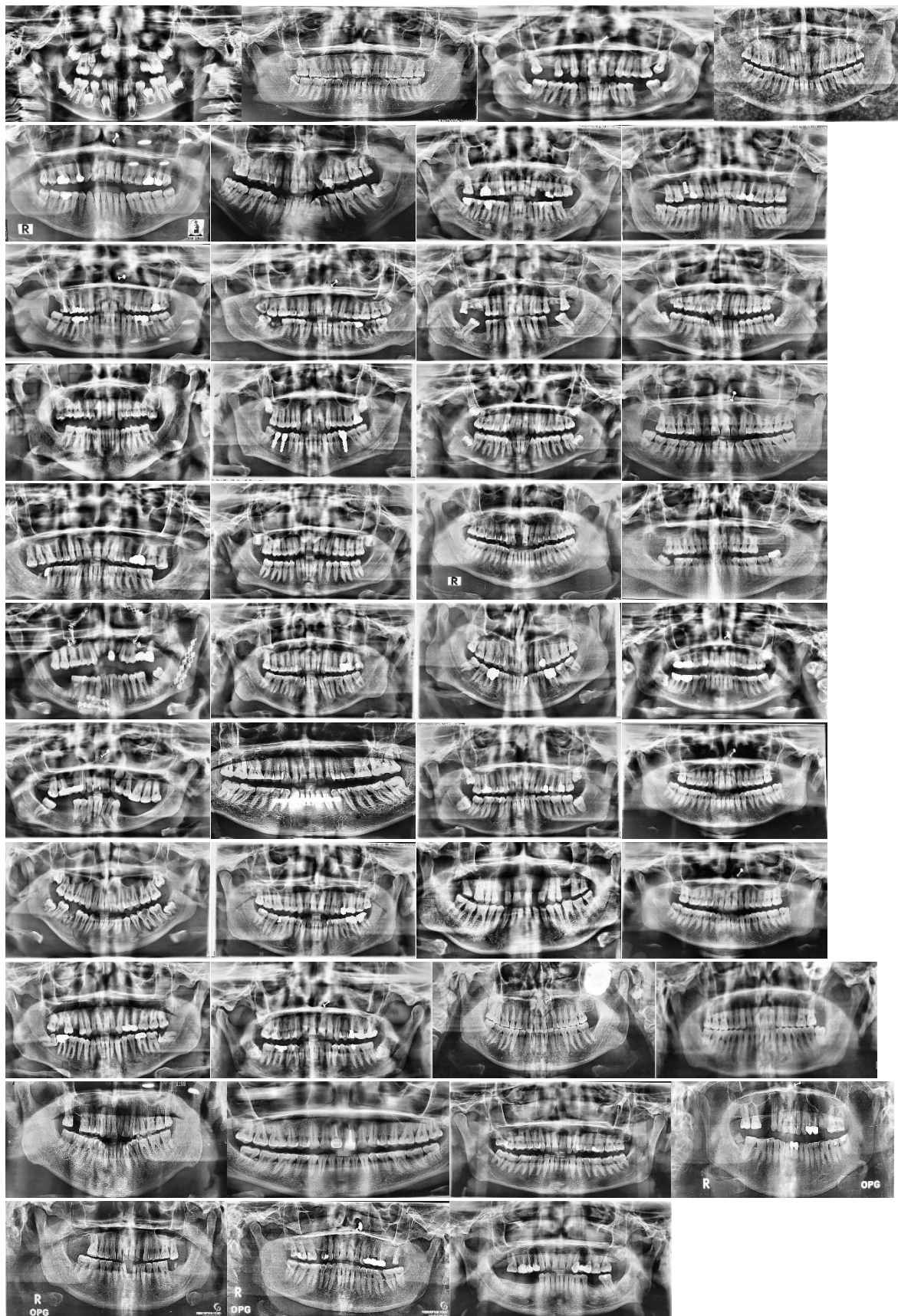


Infection:

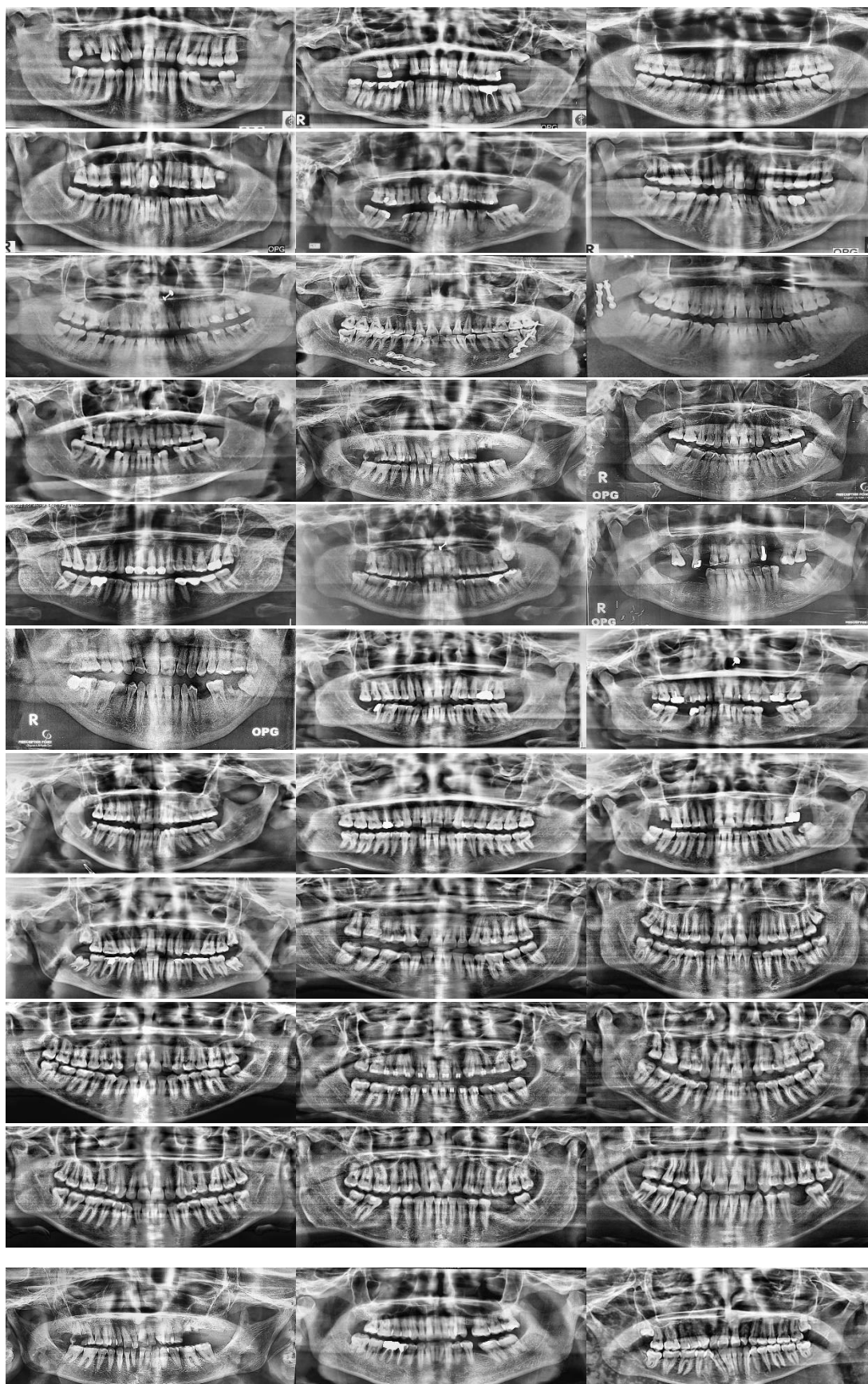


Healthy teeth:

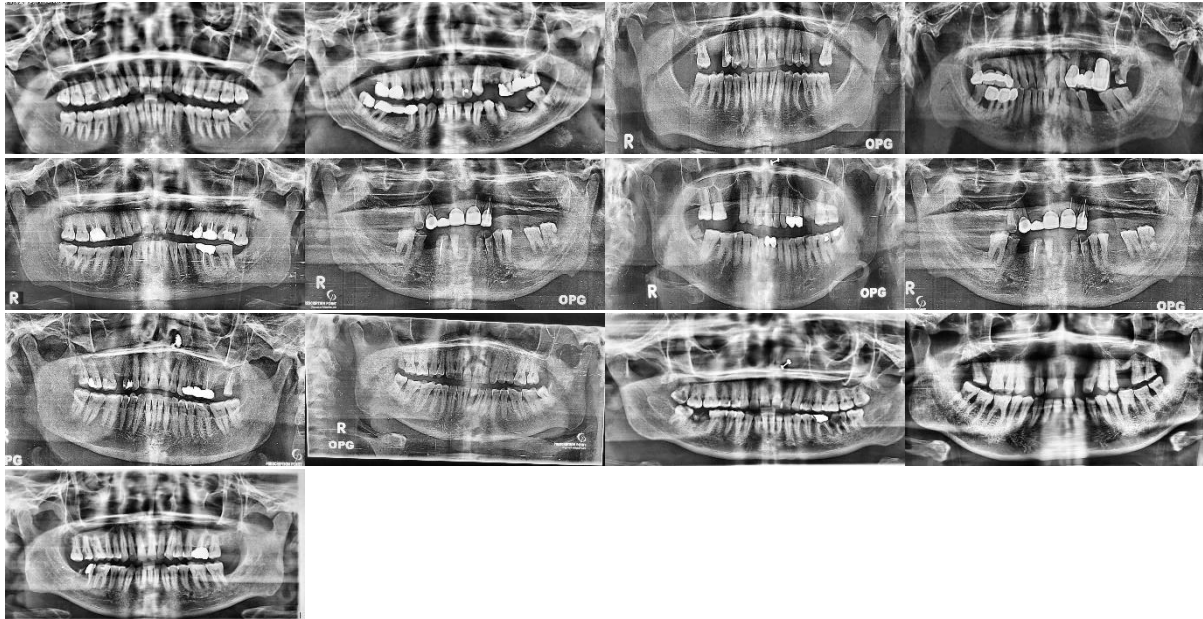




Caries:



Fractured teeth:



## PYTHON CODE IMPLEMENTATION:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plot
data=pd.read_csv('/content/dental_caries11.csv')
data.shape
data.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 52200 entries, 0 to 52199
Data columns (total 7 columns):
#   Column                                Non-Null Count  Dtype
---  -
0    County                                52200 non-null  object
1    Calendar Year                        52200 non-null  object
2    Measure                              52200 non-null  object
3    Age Filter                          52200 non-null  object
4    Users                               40533 non-null  object
5    Denominator (3 Months Continuous Eligibility) 51857 non-null  object
6    Utilization %                       40319 non-null  object
dtypes: object(7)
memory usage: 2.8+ MB
```

```
data.columns
Index(['County', 'Calendar Year', 'Measure', 'Age Filter', 'Users',
       'Denominator (3 Months Continuous Eligibility)', 'Utilization %'],
      dtype='object')
```

```
data.isnull()
```

	County	Calendar Year	Measure	Age Filter	Users	Denominator (3 Months Continuous Eligibility)	Utilization %
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...	...	...	...	...	...	...	...
52195	False	False	False	False	False	False	False
52196	False	False	False	False	False	False	False
52197	False	False	False	False	False	False	False
52198	False	False	False	False	False	False	False
52199	False	False	False	False	False	False	False

52200 rows x 7 columns

```
data.isnull().sum()
```



	0
County	0
Calendar Year	0
Measure	0
Age Filter	0
Users	11667
Denominator (3 Months Continuous Eligibility)	343
Utilization %	11881

dtype: int64

```
data.dropna()
```

	County	Calendar Year	Measure	Age Filter	Users	Denominator (3 Months Continuous Eligibility)	Utilization %
0	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age <1	164	6,878	2.38%
1	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 10-14	16,638	33,609	49.50%
2	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 1-2	4,665	16,442	28.37%
3	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 15-18	10,116	24,864	40.69%
4	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 19-20	2,678	9,457	28.32%
...	...	...	...	...	...	...	...
52195	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 21-34	1,007	7,787	12.93%
52196	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 35-44	505	4,547	11.11%
52197	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 45-64	744	6,982	10.66%
52198	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 65-74	217	2,014	10.77%
52199	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 75+	84	961	8.74%

40254 rows × 7 columns

```
data.dropna(how='any')
```

	County	Calendar Year	Measure	Age Filter	Users	Denominator (3 Months Continuous Eligibility)	Utilization %
0	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age <1	164	6,878	2.38%
1	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 10-14	16,638	33,609	49.50%
2	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 1-2	4,665	16,442	28.37%
3	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 15-18	10,116	24,864	40.69%
4	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 19-20	2,678	9,457	28.32%
...	...	...	...	...	...	...	...
52195	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 21-34	1,007	7,787	12.93%
52196	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 35-44	505	4,547	11.11%
52197	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 45-64	744	6,982	10.66%
52198	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 65-74	217	2,014	10.77%
52199	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 75+	84	961	8.74%

40254 rows × 7 columns

```
data.dropna(how='all')
```

	County	Calendar Year	Measure	Age Filter	Users	Denominator (3 Months Continuous Eligibility)	Utilization %
0	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age <1	164	6,878	2.38%
1	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 10-14	16,638	33,609	49.50%
2	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 1-2	4,665	16,442	28.37%
3	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 15-18	10,116	24,864	40.69%
4	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...	Age 19-20	2,678	9,457	28.32%
...	...	...	...	...	...	...	...
52195	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 21-34	1,007	7,787	12.93%
52196	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 35-44	505	4,547	11.11%
52197	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 45-64	744	6,982	10.66%
52198	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 65-74	217	2,014	10.77%
52199	Yuba	CY 2022	Treatment for Caries (D2000-D2999) or Caries-P...	Age 75+	84	961	8.74%

52200 rows × 7 columns

```
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion_matrix
from matplotlib.colors import ListedColormap
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
```

```

data = data.dropna()
x = data.iloc[:, [2, 3]].values
y = data.iloc[:, 4].values

import pandas as pd
from sklearn.model_selection import train_test_split
data = pd.read_csv('/content/dental_caries11.csv')
print(data.shape)
data = data.dropna(how='all', subset=data.columns[[2, 3, 4]])
print(data.shape)
x = data.iloc[:, [2, 3]].values
y = data.iloc[:, 4].values


import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
data = pd.read_csv('/content/dental_caries11.csv')
print(data.info())
print(data.columns)
data['Denominator (3 Months Continuous Eligibility)'] =
pd.to_numeric(data['Denominator (3 Months Continuous
Eligibility)'].str.replace(',', ''), errors='coerce')
data['Utilization %'] = pd.to_numeric(data['Utilization
%'].str.replace(',', ''), errors='coerce')
data['Users'] = pd.to_numeric(data['Users'].str.replace(',', ''),
errors='coerce')
for col in ['Users', 'Denominator (3 Months Continuous Eligibility)',
'Utilization %']:
    mean_val = pd.to_numeric(data[col], errors='coerce').mean()
    data[col] = data[col].fillna(mean_val)
numerical_cols = ['Denominator (3 Months Continuous Eligibility)',
'Utilization %'] # Explicitly specify columns
x = data[numerical_cols].values
y = data['Users'].values
x_train, x_test, y_train, y_test = train_test_split(x, y,
test_size=0.25, random_state=0)
print(f"Training Set (x_train):\n{x_train}")
print(f"Test Set (x_test):\n{x_test}")
st_x = StandardScaler()
x_train = st_x.fit_transform(x_train)
x_test = st_x.transform(x_test)

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 52200 entries, 0 to 52199
Data columns (total 7 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   County                52200 non-null object
 1   Calendar Year         52200 non-null object
 2   Measure               52200 non-null object
 3   Age Filter            52200 non-null object
 4   Users                 40533 non-null object
 5   Denominator (3 Months Continuous Eligibility)  51857 non-null object
 6   Utilization %         40319 non-null object
dtypes: object(7)
memory usage: 2.8+ MB
None
Index(['County', 'Calendar Year', 'Measure', 'Age Filter', 'Users',
      'Denominator (3 Months Continuous Eligibility)', 'Utilization %'],
      dtype='object')
Training Set (x_train):
[[10766.    nan]
 [ 5325.    nan]
 [13725.    nan]
 ...
 [ 1234.    nan]
 [35319.    nan]
 [50455.    nan]
Test Set (x_test):
[[ 796.    nan]
 [ 721.    nan]
 [  88.    nan]
 ...
 [ 643.    nan]
 [ 8642.    nan]
 [27681.    nan]]

```

```

import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.svm import SVC
from sklearn.metrics import confusion_matrix
import pandas as pd
data_set = pd.read_csv('/content/dental_caries11.csv')
x = data_set.iloc[:, [2, 3]].values
y = data_set.iloc[:, 4].values
x_train, x_test, y_train, y_test = train_test_split(x, y,
test_size=0.25, random_state=0)
print(f"Training Set (x_train):\n{x_train}")
print(f"Test Set (x_test):\n{x_test}")
Training Set (x_train):
[['Exams/Oral Health Evaluations (D0120, D0145, or D0150)' 'Age 15-18']
 ['Use of Sealant (D1351 or Safety Net Clinics 03 Encounters with ICD 10 Code Z98810)'
 'Age 45-64']
 ['Use of Sealant (D1351 or Safety Net Clinics 03 Encounters with ICD 10 Code Z98810)'
 'Age 10-14']
 ...
 ['Treatment for Caries (D2000-D2999) or Caries-Preventive Procedure (D1203-D1208, D1310, D1330, D1351 or Safety Net Clinics 03 Encounters with identified ICD 10 Codes)'
 'Age 1-2']
 ['Annual Dental Visit (D0100 - D9999, CPT 99188 or Safety Net Clinics 03 Encounters)'
 'Age 21-34']
 ['Use of Preventive Services (D1000 - D1999)' 'Age 45-64']]
Test Set (x_test):
[['Annual Dental Visit (D0100 - D9999 or Safety Net Clinics 03 Encounters)'
 'Age <1']
 ['Use of Dental Treatment Services (D2000 - D9999)' 'Age 45-64']
 ['Exams/Oral Health Evaluations (D0120, D0145, or D0150 or Safety Net Clinics 03 Encounters with identified ICD 10 Codes)'
 'Age <1']
 ...
 ['Use of Preventive Services (D1000 - D1999 or Safety Net Clinics 03 Encounters with identified ICD 10 Codes)'
 'Age 10-14']
 ['Use of Sealant (D1351 or Safety Net Clinics 03 Encounters with ICD 10 Code Z98810)'
 'Age 75+']
 ['Use of Restorative Services (D2000 - D2999)' 'Age 10-14']]

```

```

import pandas as pd
import numpy as np
from sklearn.metrics import multilabel_confusion_matrix
csv_file_path = '/content/dental_caries11.csv'
df = pd.read_csv(csv_file_path)
df['Users'] = df['Users'].replace('nan', np.nan)
def convert_to_numeric(x):
    try:
        return int(x.replace(',', ''))
    except (ValueError, TypeError):

```

```

        return np.nan
df["Users"] = df["Users"].astype(str).apply(convert_to_numeric)
df = df.dropna(subset=['Users'])
y_true = np.array(df["Users"].tolist())
np.random.seed(42)
y_pred = np.random.randint(0, 2, size=y_true.shape)
conf_matrices = multilabel_confusion_matrix(y_true, y_pred)
label_names = ["Denominator (3 Months Continuous Eligibility)",
               "Utilization %", "Users"]
conf_matrices_dict = {label: conf_matrices[i] for i, label in
                      enumerate(label_names)}
for label, matrix in conf_matrices_dict.items():
    print(f"Confusion Matrix for {label}:")
    print(matrix)
    print()
Confusion Matrix for Denominator (3 Months Continuous Eligibility):
[[19350 19331]
 [ 948  904]]

Confusion Matrix for Utilization %:
[[20235 20298]
 [    0    0]]

Confusion Matrix for Users:
[[40381    0]
 [  152    0]]

```

```

from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error
import pandas as pd
data = pd.read_csv('/content/dental_caries11.csv')
X = data[['Denominator (3 Months Continuous Eligibility)', 'Users']]
y = data['Utilization %']
X['Denominator (3 Months Continuous Eligibility)'] =
pd.to_numeric(X['Denominator (3 Months Continuous
Eligibility)'].str.replace(',', ''), errors='coerce')
X['Users'] = pd.to_numeric(X['Users'].str.replace(',', ''),
errors='coerce')
y = pd.to_numeric(y.str.rstrip('%'), errors='coerce')
X = X.fillna(X.mean())
y = y.fillna(y.mean())
X = pd.get_dummies(X, drop_first=True)
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.3, random_state=42)
model = RandomForestRegressor(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')

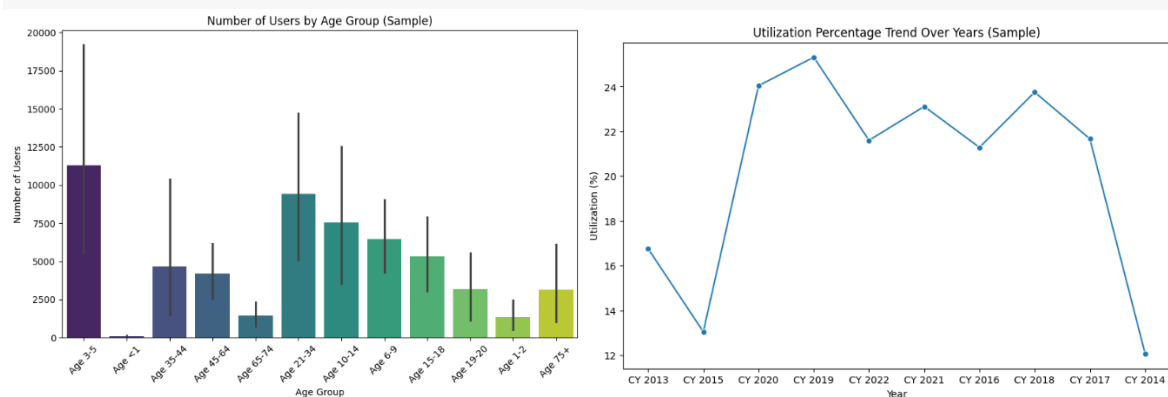
import pandas as pd
import seaborn as sns

```

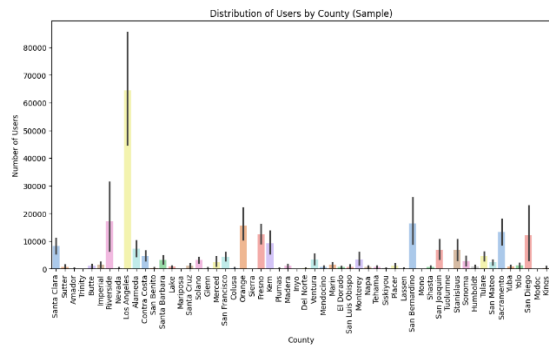
```

import matplotlib.pyplot as plt
file_path = "dental_caries11.csv"
dental_data = pd.read_csv(file_path)
dental_data = dental_data.drop(columns=['Unnamed: 0'], errors='ignore')
dental_data['Users'] = dental_data['Users'].str.replace(',',
').astype(float)
dental_data['Denominator (3 Months Continuous Eligibility)'] =
dental_data['Denominator (3 Months Continuous
Eligibility)'].str.replace(',', ' ').astype(float)
dental_data['Utilization %'] = dental_data['Utilization
%'].str.replace('%', ' ').astype(float)
dental_data = dental_data.dropna()
dental_data_sample = dental_data.sample(n=500, random_state=42)
plt.figure(figsize=(10, 6))
sns.barplot(x='Age Filter', y='Users', data=dental_data_sample,
palette='viridis')
plt.title("Number of Users by Age Group (Sample)")
plt.xlabel("Age Group")
plt.ylabel("Number of Users")
plt.xticks(rotation=45)
plt.show()
plt.figure(figsize=(10, 6))
sns.lineplot(x='Calendar Year', y='Utilization %',
data=dental_data_sample, marker='o', ci=None)
plt.title("Utilization Percentage Trend Over Years (Sample)")
plt.xlabel("Year")
plt.ylabel("Utilization (%)")
plt.show()
plt.figure(figsize=(12, 6))
sns.barplot(x='County', y='Users', data=dental_data_sample,
palette='pastel')
plt.title("Distribution of Users by County (Sample)")
plt.xlabel("County")
plt.ylabel("Number of Users")
plt.xticks(rotation=90)
plt.show()

```





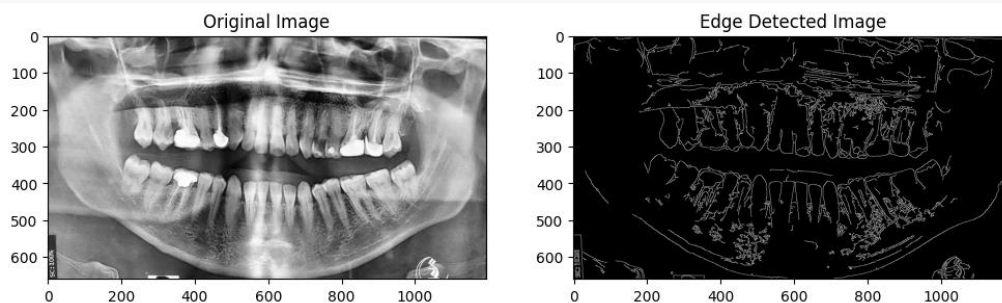


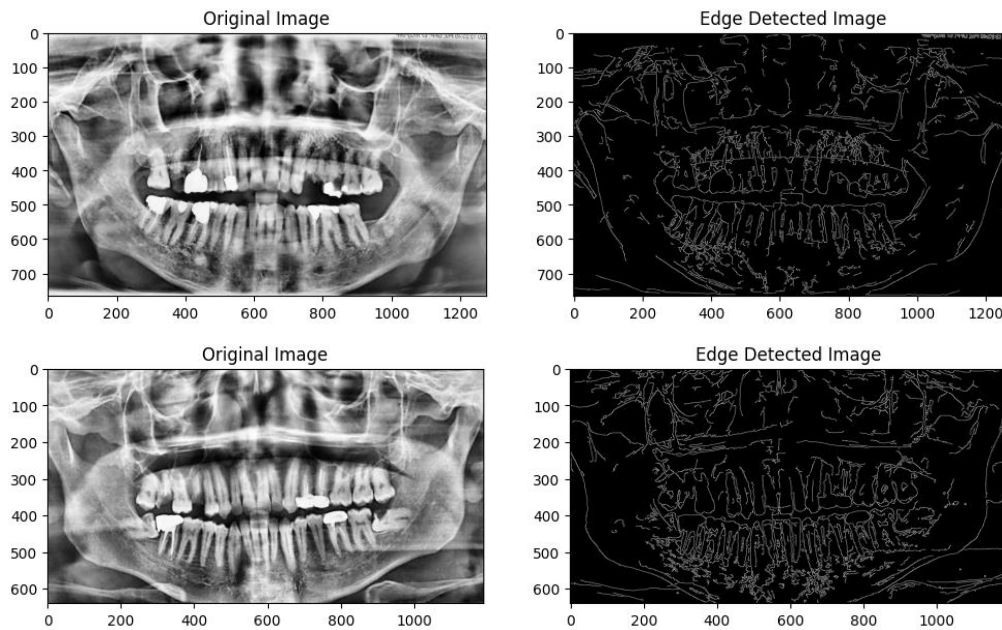
```
import cv2
import numpy as np
import matplotlib.pyplot as plt

image_paths = [
    "/content/53.jpg",
    "/content/54.jpg",
    "/content/6.jpg",
    "/content/59.jpg",
    "/content/8.jpg"
]

def process_and_display_images(image_paths):
    for img_path in image_paths:
        image = cv2.imread(img_path, cv2.IMREAD_GRAYSCALE)
        blurred = cv2.GaussianBlur(image, (5, 5), 0)
        edges = cv2.Canny(blurred, 50, 150)
        plt.figure(figsize=(12, 6))
        plt.subplot(1, 2, 1)
        plt.title("Original Image")
        plt.imshow(image, cmap='gray')
        plt.subplot(1, 2, 2)
        plt.title("Edge Detected Image")
        plt.imshow(edges, cmap='gray')
        plt.show()

process_and_display_images(image_paths)
```





```
from sklearn.metrics import mean_squared_error, r2_score
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
r2 = r2_score(y_test, y_pred)
print(f'R2 Score: {r2}')
```

**Mean Squared Error: 0.40017244446028927**  
**R<sup>2</sup> Score: 0.9978955900440231**

```
from sklearn.metrics import mean_squared_error, r2_score
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
r2 = r2_score(y_test, y_pred)
print(f'R2 Score: {r2}')
```

**Mean Squared Error: 0.40017244446028927**  
**R<sup>2</sup> Score: 0.9978955900440231**

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
from sklearn.preprocessing import LabelEncoder
file_path = "/content/dental_caries11.csv"
data = pd.read_csv(file_path)
print("Dataset Info:")
print(data.info())
print("\nFirst 5 rows of the dataset:")
```

```

print(data.head())
X = data.drop(columns=["Users"])
y = data["Users"]
X.dropna(inplace=True)
y = y.loc[X.index]
print("\nMissing values after dropping:")
print(X.isnull().sum())
for col in X.select_dtypes(include=['object']).columns:
    le = LabelEncoder()
    X[col] = le.fit_transform(X[col])
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
clf = RandomForestClassifier(n_estimators=100, random_state=42)
clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"\n✓ Model Accuracy: {accuracy * 100:.2f}%")
print("\nClassification Report:")
print(classification_report(y_test, y_pred))

```

```

data.columns
#      Column                                     Non-Null Count  Dtype
---  -
0      County                                     52200 non-null      object
1      Calendar Year                             52200 non-null      object
2      Measure                                     52200 non-null      object
3      Age Filter                                  52200 non-null      object
4      Users                                       40533 non-null      object
5      Denominator (3 Months Continuous Eligibility)  51857 non-null      object
6      Utilization %                             40319 non-null      object
dtypes: object(7)
memory usage: 2.8+ MB
None

```

First 5 rows of the dataset:

	County	Calendar Year	Measure \
0	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...
1	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...
2	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...
3	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...
4	Alameda	CY 2013	Annual Dental Visit (D0100 - D9999 or Safety N...

	Age Filter	Users	Denominator (3 Months Continuous Eligibility) \
0	Age <1	164	6,878
1	Age 10-14	16,638	33,609
2	Age 1-2	4,665	16,442
3	Age 15-18	10,116	24,864
4	Age 19-20	2,678	9,457

	Utilization %
0	2.38%
1	49.50%
2	28.37%
3	40.69%
4	28.32%