

Data context: COVID-19 hospitalizations (ICU or Med/Surg), mortality by age groups and tests over time.

Database scheme:



Executive Summary

The COVID-19 pandemic has significantly strained healthcare systems worldwide, including San Francisco, where hospitals faced an unprecedented surge in ICU (intensive care) and Med/Surg (acute care) patient counts. This report examines hospitalization trends, mortality data, and testing outcomes across various years to identify correlations and inform future preparedness. Key findings include notable spikes in ICU admissions during January 2021, the high vulnerability of older populations (80+ age group), and the correlation between rising positive test counts and increased hospitalizations and mortality. These insights underscore the critical need for scalable healthcare infrastructure, targeted interventions for vulnerable populations, and robust testing programs to mitigate future healthcare crises.

Objectives

1. To analyze hospitalization trends in ICU and Med/Surg departments in San Francisco hospitals.
2. To explore mortality patterns across different age groups and years.
3. To correlate hospitalization rates, mortality trends, and test outcomes to inform healthcare strategies.

Introduction

The COVID-19 pandemic disrupted global healthcare systems, exposing vulnerabilities and challenging hospital capacities. In San Francisco, like many other urban centers, the pressure on ICU and Med/Surg departments highlighted the critical need for data-driven decision-making. This report evaluates hospitalization and mortality trends over five years (2020–2024) to understand the relationship between testing outcomes, patient counts, and mortality rates, focusing on periods of significant changes.

Problem Definition: COVID-19 Impact on Hospitalization and Mortality in San Francisco

The COVID-19 pandemic significantly impacted health systems globally, and understanding how it affected hospitalizations and mortality in a specific region is crucial for future preparedness. In San Francisco, as with many urban centres, the

pressure on healthcare services, especially in ICU and Med/Surg departments, led to a rapidly evolving situation. This analysis aims to explore how hospitalization trends and mortality patterns evolved in San Francisco, with a focus on demographic variations and the factors that influenced these trends.

Overarching Questions

1. How did COVID-19 hospitalization trends evolve in San Francisco, and what factors influenced patient counts across ICU and Med/Surg departments?
2. What were the mortality patterns for COVID-19 patients in San Francisco hospitals, and how did these vary by demographic group over time?

Related Sub-Problems

- **Hospitalization Data Analysis**

- What trends emerged in COVID-19 patient counts within ICU and Med/Surg departments at San Francisco hospitals?

We want to compare the trends of the patient count across the years for patients that were in intensive care (ICU) and acute care (Med/Surg), for COVID-positive patients.

Results:

Year	COVID_Patient_count_per_Year	Year	COVID_Patient_count_per_Year_MED
2020	6456	2020	14012
2021	6863	2021	16370
2022	4953	2022	28721
2023	2379	2023	17278
2024	498	2024	4656

Comments: We can observe from the results that the number of patients in acute care was higher than patients in intensive care. 2021 was the year with the greatest number of patients in intensive care, while 2022 was the year with the greatest number of patients in acute care. Despite that the number of patients in acute care was increasing, the number of patients in a critical situation was not.

Queries:

```
SELECT dates.Year, SUM(ICU.patient_count) AS COVID_Patient_count_per_Year
FROM archived_covid_19_icu AS ICU
LEFT JOIN date_table AS dates
ON ICU.Date_id = dates.Date_id
WHERE ICU.covid_status = "COVID+"
GROUP BY dates.Year ;
```

```
SELECT dates.Year, SUM(MED.patient_count) AS COVID_Patient_count_per_Year_MED
FROM archive_covid_19_med_sug AS MED
LEFT JOIN date_table AS dates
ON MED.Date_id = dates.Date_id
WHERE MED.covid_status = "COVID+"
```

GROUP BY dates.Year ;

🚩 Are there identifiable periods with significant changes in hospitalizations for patients in intensive care (ICU)?

We will focus on year 2021 since that was the year with the greatest number of patients hospitalized.

Results:

Month	COVID_Patient_count_per_Month	Month	Positive_tests	Negative_tests
Jan	1720	Jan	9095	215846
Feb	956	Feb	3219	168215
Mar	357	Mar	1505	172113
Apr	207	Apr	1284	150824
May	157	May	681	118760
Jun	143	Jun	603	91631
Jul	417	Jul	5860	119853
Aug	1125	Aug	7826	196658
Sep	695	Sep	3893	195064
Oct	463	Oct	2366	157327
Nov	308	Nov	2481	145156
Dec	315	Dec	20282	219181

ICU_ratio

0.1161

Comments: From the results we can observe that January is the month with the greatest number of ICU hospitalizations and the number goes down until August when there is a rebound two months after the lowest number of hospitalizations during the year.

We checked the number of positive and negative tests for the same period.

We can see on that the number of positive tests has the same behaviour as the number of patients hospitalized in ICU.

And from all positive tests during 2021, 11.6% of those people needed ICU.

To summarize this part, based on the spikes in ICU admissions noted in January 2021, hospitals should establish plans for quickly expanding ICU capacity during critical periods. This could involve converting Med/Surg beds to ICU temporarily or setting up additional ICU units.

Since 2022 showed an increase in Med/Surg hospitalizations, it's essential to ensure that acute care units can handle a high patient load without requiring ICU care. Invest in scalable infrastructure that allows for a rapid shift between care levels.

Queries:

```
SELECT dates.Month AS Month, SUM(ICU.patient_count) AS COVID_Patient_count_per_Month
FROM archived_covid_19_icu AS ICU
LEFT JOIN date_table AS dates
ON ICU.Date_id = dates.Date_id
WHERE ICU.covid_status = 'COVID+' AND dates.Year = 2021
GROUP BY dates.Month
ORDER BY dates.Month_Number ;
```

```

SELECT dates.Month, SUM(covid.pos) AS Positive_tests, SUM(covid.neg) AS Negative_tests
FROM covid_19_overtime AS covid
LEFT JOIN date_table AS dates
ON covid.Date_id = dates.Date_id
WHERE dates.Year = 2021
GROUP BY dates.Month
ORDER BY dates.Month_Number ;

```

```

SELECT ICU.COVID_Patient_count_per_Month/tests.Positive_tests AS ICU_ratio
FROM
(SELECT dates.Month AS Month, SUM(ICU.patient_count) AS COVID_Patient_count_per_Month
FROM archived_covid_19_icu AS ICU
LEFT JOIN date_table AS dates
ON ICU.Date_id = dates.Date_id
WHERE ICU.covid_status = 'COVID+' AND dates.Year = 2021
) AS ICU
INNER JOIN
(SELECT dates.Month AS Month, SUM(covid.pos) AS Positive_tests, SUM(covid.neg) AS Negative_tests
FROM covid_19_overtime AS covid
LEFT JOIN date_table AS dates
ON covid.Date_id = dates.Date_id
WHERE dates.Year = 2021
) AS tests
ON ICU.Month = tests.Month;

```

- **Mortality Data Analysis in San Francisco Hospitals**

- 🚩 How did COVID-19-related deaths among patients vary across age groups and other demographic characteristics in San Francisco hospitals?

Results:

age_group	total_deaths ▾ 1
80+	778
70-79	228
60-69	165
50-59	86
21-24	3
0-4	0
18-20	0

Comments: We can observe that older groups had more deaths.

Queries:

SELECT

characteristic_group AS age_group,

SUM(new_deaths) AS total_deaths

FROM

combined_mortality

GROUP BY

characteristic_group

ORDER BY

total_deaths DESC;

CREATE TABLE combined_mortality AS

SELECT * FROM 0_4

UNION ALL

SELECT * FROM 18_20

UNION ALL

SELECT * FROM 21_24

UNION ALL

SELECT * FROM 50_59

UNION ALL

SELECT * FROM 60_69

UNION ALL

SELECT * FROM 60_69

UNION ALL

SELECT * FROM 70_79

UNION ALL

SELECT * FROM 80_

;

What patterns can be observed in the progression of cumulative COVID-19 deaths?

Results:

age_group	total_cumulative_deaths ▾ 1
80+	16186
60-69	3866
70-79	3576
50-59	1593
21-24	312
0-4	0
18-20	0

Comments:

The distribution of total cumulative COVID-19 deaths across age groups in 2020 highlights a significant disparity in mortality rates. The 80+ age group experienced the highest number of cumulative deaths, with 16,186 deaths, underscoring their vulnerability to severe outcomes of the virus. This is followed by the 60-69 (3,866 deaths) and 70-79 (3,576 deaths) age groups, indicating elevated risks among older adults. In contrast, younger age groups, such as 21-24 and below, recorded minimal or no cumulative deaths, as seen with 312 deaths in the 21-24 age group and zero deaths in the 0-4 and 18-20 groups. These patterns emphasize the critical need for targeted protective measures for older populations to mitigate the mortality burden of the pandemic.

Queries:

SELECT

characteristic_group AS age_group,

SUM(cumulative_deaths) AS total_cumulative_deaths

FROM

combined_mortality

WHERE

YEAR = 2020

GROUP BY

characteristic_group

ORDER BY

total_cumulative_deaths DESC;

🚦 Were there specific periods where mortality rates spiked, and do these correspond to shifts in hospitalization data?

Results:

Year	Month	Total_New_Deaths	Deaths_Spike	Total_Patients
2020	Dec	395	295	11083028
2021	Jan	545	295	12281449

Comment: The table result highlights a significant observation in the analysis of mortality data during the COVID-19 pandemic. In December 2020, the total number of new deaths was 395, with a noticeable spike of 295 deaths compared to previous months. This trend continued into January 2021, where new deaths increased further to 545, maintaining the same spike of 295. This indicates a critical period of increased mortality, potentially corresponding to heightened COVID-19 cases or hospitalizations during this time. The accompanying increase in total patients, from approximately 11 million in December 2020 to over 12 million in January 2021, further supports the notion of increased healthcare burden during this period. This correlation between rising patient numbers and mortality spikes could indicate healthcare system strain or the impact of worsening pandemic conditions.

Queries:

WITH CombinedData AS (

SELECT

dt.Date_id,

dt.Date,

dt.Month,

dt.Year,

dt.Day_of_week,

'MedSurg' AS Hospital_Type,

ms.patient_count,

cm.new_deaths,

cm.cumulative_deaths

FROM date_table dt

JOIN archive_covid_19_med_sug ms ON dt.Date_id = ms.Date_id

LEFT JOIN combined_mortality cm ON dt.Date_id = cm.Date_id

UNION ALL

SELECT

dt.Date_id,

dt.Date,

dt.Month,

dt.Year,

dt.Day_of_week,

'ICU' AS Hospital_Type,

icu.patient_count,

```

        cm.new_deaths,
        cm.cumulative_deaths
FROM date_table dt
JOIN archived_covid_19_icu icu ON dt.Date_id = icu.Date_id
LEFT JOIN combined_mortality cm ON dt.Date_id = cm.Date_id
UNION ALL
SELECT
    dt.Date_id,
    dt.Date,
    dt.Month,
    dt.Year,
    dt.Day_of_week,
    'AgeGroup' AS Hospital_Type,
    cm.population_estimate AS patient_count,
    cm.new_deaths,
    cm.cumulative_deaths
FROM date_table dt
JOIN combined_mortality cm ON dt.Date_id = cm.Date_id
)
, MonthlyData AS (
    SELECT
        Year,
        Month,
        SUM(patient_count) AS Total_Patients,
        SUM(new_deaths) AS Total_New_Deaths,
        MAX(cumulative_deaths) AS Max_Cumulative_Deaths
    FROM CombinedData
    GROUP BY Year, Month
)
, MortalitySpikes AS (
    SELECT
        Year,
        Month,
        Total_New_Deaths,
        LAG(Total_New_Deaths) OVER (ORDER BY Year, Month) AS Previous_Deaths,

```



```

        Total_New_Deaths - LAG(Total_New_Deaths) OVER (ORDER BY Year, Month) AS Deaths_Spike
    FROM MonthlyData
)
SELECT
    ms.Year,
    ms.Month,
    ms.Total_New_Deaths,
    ms.Deaths_Spike,
    md.Total_Patients
FROM MortalitySpikes ms
LEFT JOIN MonthlyData md
    ON ms.Year = md.Year AND ms.Month = md.Month
WHERE ms.Deaths_Spike > (SELECT AVG(Deaths_Spike) + 2 * STDDEV(Deaths_Spike) FROM MortalitySpikes)
ORDER BY ms.Year, ms.Month;

```

Results:

Year	Total_Patients	Total_New_Deaths	Total_Positive_Tests	Total_Negative_Tests
2020	111053522	1149	976752	36236523
2021	144214424	2065	2068325	68271980
2022	144296107	2165	5416600	47837405
2023	144184429	537	590366	9440039
2024	67115026	68	67399	2324462

Comments: The table presents an integrated yearly analysis of hospitalization and mortality trends from 2020 to 2024. It shows a steady increase in total patients from 2020 to 2023, with a significant drop in 2024. The total new deaths peaked in 2022, aligning with the highest numbers of positive tests in the same year, which suggests a correlation between increased positive cases and mortality rates. Conversely, the total negative tests were highest in 2023, indicating improved testing efforts, though the deaths significantly decreased that year. This trend highlights potential improvements in healthcare management and COVID-19 response over time, particularly after 2022.

Queries:

```

WITH CombinedData AS (
    SELECT
        dt.Date_id,
        dt.Date,
        dt.Year,
        dt.Day_of_week,
        'MedSurg' AS Hospital_Type,
        ms.patient_count,
        cm.new_deaths,
        cm.cumulative_deaths,

```

```

co.pos AS positive_tests,

co.neg AS negative_tests

FROM date_table dt

JOIN archive_covid_19_med_sug ms ON dt.Date_id = ms.Date_id

LEFT JOIN combined_mortality cm ON dt.Date_id = cm.Date_id

LEFT JOIN covid_19_overtime co ON dt.Date_id = co.Date_id

UNION ALL

SELECT

```

```

dt.Date_id,

dt.Date,

dt.Year,

dt.Day_of_week,

'ICU' AS Hospital_Type,

icu.patient_count,

cm.new_deaths,

cm.cumulative_deaths,

co.pos AS positive_tests,

co.neg AS negative_tests

```

```

FROM date_table dt

JOIN archived_covid_19_icu icu ON dt.Date_id = icu.Date_id

LEFT JOIN combined_mortality cm ON dt.Date_id = cm.Date_id

LEFT JOIN covid_19_overtime co ON dt.Date_id = co.Date_id

UNION ALL

SELECT

```

```

dt.Date_id,

dt.Date,

dt.Year,

dt.Day_of_week,

'AgeGroup' AS Hospital_Type,

cm.population_estimate AS patient_count,

cm.new_deaths,

cm.cumulative_deaths,

co.pos AS positive_tests,

co.neg AS negative_tests

```

```

FROM date_table dt

```

```

JOIN combined_mortality cm ON dt.Date_id = cm.Date_id

LEFT JOIN covid_19_overtime co ON dt.Date_id = co.Date_id
)
, YearlyData AS (
SELECT
    Year,
    SUM(patient_count) AS Total_Patients,
    SUM(new_deaths) AS Total_New_Deaths,
    MAX(cumulative_deaths) AS Max_Cumulative_Deaths,
    SUM(positive_tests) AS Total_Positive_Tests,
    SUM(negative_tests) AS Total_Negative_Tests
FROM CombinedData
GROUP BY Year
)
SELECT
    Year,
    Total_Patients,
    Total_New_Deaths,
    Total_Positive_Tests,
    Total_Negative_Tests
FROM YearlyData
ORDER BY Year;

```

- **Integrated Analysis of Hospitalization and Mortality**

🔗 Can relationships between ICU and Med/Surg hospitalization rates and mortality trends be identified?

Results:

	Year	Total_MedSurg_Hospitalizations	Total_ICU_Hospitalizations	Total_New_Deaths	Max_Cumulative_Deaths
<input type="checkbox"/> Edit Copy Delete	2020	118153	46587	904	159
<input type="checkbox"/> Edit Copy Delete	2021	127386	50008	1652	378
<input type="checkbox"/> Edit Copy Delete	2022	221963	37114	1732	652
<input type="checkbox"/> Edit Copy Delete	2023	129787	17612	398	753
<input type="checkbox"/> Edit Copy Delete	2024	26526	2760	38	777

Comments: The table provides an integrated yearly analysis of hospitalizations (Med/Surg and ICU) and mortality trends from 2020 to 2024. It highlights the total number of hospitalizations in Med/Surg and ICU facilities alongside total new deaths and the maximum cumulative deaths recorded each year. The data shows that 2022 had the highest Med/Surg hospitalizations (221,963) and ICU hospitalizations (37,114), correlating with a significant number of new deaths (1,732) and cumulative deaths (652). Conversely, 2024 shows a sharp decline in hospitalizations and new deaths, indicating a reduction in severe COVID-19 cases or improved management over time. These trends suggest a relationship between hospitalization rates and mortality, with years of high hospitalizations typically aligning with higher mortality figures.

Queries:

WITH HospitalizationData AS (

SELECT

dt.Date_id,

dt.Date,

dt.Year,

'MedSurg' AS Hospital_Type,

ms.patient_count AS MedSurg_Hospitalizations,

0 AS ICU_Hospitalizations

FROM date_table dt

JOIN archive_covid_19_med_sug ms ON dt.Date_id = ms.Date_id

UNION ALL

SELECT

dt.Date_id,

dt.Date,

dt.Year,

'ICU' AS Hospital_Type,

0 AS MedSurg_Hospitalizations,

icu.patient_count AS ICU_Hospitalizations

FROM date_table dt

JOIN archived_covid_19_icu icu ON dt.Date_id = icu.Date_id

),

MortalityData AS (

SELECT

dt.Date_id,

dt.Date,

dt.Year,

cm.new_deaths,

cm.cumulative_deaths

FROM date_table dt

JOIN combined_mortality cm ON dt.Date_id = cm.Date_id

),

CombinedAnalysis AS (

SELECT

h.Year,

```

SUM(h.MedSurg_Hospitalizations) AS Total_MedSurg_Hospitalizations,

SUM(h.ICU_Hospitalizations) AS Total_ICU_Hospitalizations,

SUM(m.new_deaths) AS Total_New_Deaths,

MAX(m.cumulative_deaths) AS Max_Cumulative_Deaths

FROM HospitalizationData h

LEFT JOIN MortalityData m ON h.Date_id = m.Date_id

GROUP BY h.Year

)

SELECT

Year,

Total_MedSurg_Hospitalizations,

Total_ICU_Hospitalizations,

Total_New_Deaths,

Max_Cumulative_Deaths
















FROM CombinedAnalysis

ORDER BY Year;

```

🔗 How do changes in patient count within San Francisco hospitals correlate with mortality trends over time?

Results:

[-] T [-]				▼ Year	Total_Hospitalizations	Total_Positive_Tests	Total_New_Deaths	Max_Cumulative_Deaths
<input type="checkbox"/>	 Edit	 Copy	 Delete	2020	25810	30775	245	159
<input type="checkbox"/>	 Edit	 Copy	 Delete	2021	25342	59095	413	378
<input type="checkbox"/>	 Edit	 Copy	 Delete	2022	37011	154760	433	652
<input type="checkbox"/>	 Edit	 Copy	 Delete	2023	21057	21158	139	753
<input type="checkbox"/>	 Edit	 Copy	 Delete	2024	5154	9670	30	778

Comments: The table illustrates the yearly relationship between total hospitalizations, positive tests, new deaths, and cumulative deaths, providing insight into the correlation between these metrics. In 2020 and 2021, a clear pattern emerges where an increase in positive tests correlates with higher hospitalizations and new deaths, indicating the direct impact of rising COVID-19 cases on hospital burden and mortality. By 2022, while positive tests surged significantly, hospitalizations and deaths showed a relatively smaller increase, potentially reflecting improved medical responses or vaccine impacts. In 2023 and 2024, both positive tests and hospitalizations decreased sharply, and new deaths followed the trend, though cumulative deaths continue to rise slightly due to the residual impact of earlier periods. This suggests that while case management and preventive measures have improved, the lingering effects of severe infections still contribute to overall mortality trends.

Queries:

Code

```

WITH HospitalizationData AS (

SELECT

dt.Year,

SUM(ms.patient_count) AS Total_MedSurg_Hospitalizations,

0 AS Total_ICU_Hospitalizations

```

```

FROM date_table dt

JOIN archive_covid_19_med_sug ms ON dt.Date_id = ms.Date_id

GROUP BY dt.Year

UNION ALL

SELECT

    dt.Year,

    0 AS Total_MedSurg_Hospitalizations,

    SUM(icu.patient_count) AS Total_ICU_Hospitalizations

FROM date_table dt

JOIN archived_covid_19_icu icu ON dt.Date_id = icu.Date_id

GROUP BY dt.Year

),

MortalityData AS (

    SELECT

        dt.Year,

        SUM(cm.new_deaths) AS Total_New_Deaths,

        MAX(cm.cumulative_deaths) AS Max_Cumulative_Deaths

    FROM date_table dt

    JOIN combined_mortality cm ON dt.Date_id = cm.Date_id

    GROUP BY dt.Year

),

TestingData AS (

    SELECT

        dt.Year,

        SUM(c19o.pos) AS Total_Positive_Tests

    FROM date_table dt

    JOIN covid_19_overtime c19o ON dt.Date_id = c19o.Date_id

    GROUP BY dt.Year

),

CombinedData AS (

    SELECT

        h.Year,

        SUM(h.Total_MedSurg_Hospitalizations) AS Total_MedSurg_Hospitalizations,

        SUM(h.Total_ICU_Hospitalizations) AS Total_ICU_Hospitalizations,

        m.Total_New_Deaths,

```

```

        m.Max_Cumulative_Deaths,
        t.Total_Positive_Tests
    FROM HospitalizationData h
    LEFT JOIN MortalityData m ON h.Year = m.Year
    LEFT JOIN TestingData t ON h.Year = t.Year

    GROUP BY h.Year, m.Total_New_Deaths, m.Max_Cumulative_Deaths, t.Total_Positive_Tests
)

SELECT
    Year,
    Total_MedSurg_Hospitalizations + Total_ICU_Hospitalizations AS Total_Hospitalizations,
    Total_Positive_Tests,
    Total_New_Deaths,
    Max_Cumulative_Deaths
FROM CombinedData
ORDER BY Year;

```

Conclusion

This analysis highlights the dynamic interplay between hospitalization rates, testing outcomes, and mortality trends during the COVID-19 pandemic in San Francisco. The surge in ICU and Med/Surg hospitalizations during 2021 and 2022 aligned closely with increases in positive tests and new deaths. However, the later decline in hospitalizations and deaths suggests improved public health measures and vaccine effectiveness. Notably, older age groups were disproportionately affected, emphasizing the need for targeted interventions for vulnerable populations.

Recommendations

Scalable Healthcare Infrastructure: Hospitals should develop flexible capacity plans, enabling rapid conversion of Med/Surg beds to ICU units during critical periods.

Targeted Interventions: Focus on protecting high-risk groups, particularly individuals aged 60 and above, through vaccination, booster programs, and tailored healthcare strategies.

Expanded Testing and Monitoring: Continue investing in large-scale testing and surveillance to detect trends early and implement timely interventions.

Data-Driven Decision-Making: Regularly update predictive models based on hospitalization and mortality data to anticipate healthcare needs during potential future waves.

Public Awareness Campaigns: Educate communities on preventive measures and the importance of early testing and treatment to reduce severe outcomes.