Datathon Meir Hospital

Authors:
Eden Giladi 209514694
Linoy Halifa 201052198
Hadar Mentel 205642366

GitHub repository link: HadarMentel/Datathon-Meir-hospital

Table of Contents:

1. Background, Objectives, and Problem Definition

1.1 General Background

The datathon took place at Afeka College and was based on real data provided by Meir Medical Center. The main objective was to leverage historical hospital data to improve the planning and management of operational workloads across various departments.

Hospitals today face significant challenges in staffing allocation and workload management. Unexpected surges in demand can lead to understaffing during critical hours, while quieter periods may result in inefficient resource utilization.

1.2 Project Goals

- Analyze historical request data from different departments at Meir Medical Center (e.g., blood tests, cleaning, doctor visits, etc.).
- Develop a reliable forecasting model to support hospital operations in proactively planning staff allocation based on predicted workload by hour and day.
- Establish a foundation for a smart system that can recommend and automate staff scheduling in alignment with forecasted demand.

1.3 Defined Problem

- There is a considerable mismatch between current staff allocation and actual workload.
- Heavy workload peaks were identified at unexpected times, particularly during late-night and early morning hours.
- There is a need for a more data-driven, predictive approach to workforce planning, moving away from fixed shift patterns and reliance on historical experience alone.

1.4 Department Categorization

To better understand workload patterns, departments were divided into two main categories:

- Inpatient Departments: Departments A, B, C, and Trauma.
- Outpatient Departments: All other departments within the hospital.

1.5 Data Description

1.5.1 Data Source

The dataset used for this project was provided by Meir Medical Center. It was delivered in a CSV format and contained detailed records of service requests (tasks) performed in the hospital during a specified time period. The modeling and analysis focused on data starting from July 1, 2024, as instructed.

1.5.2 Dataset Structure

The data file included the following columns:

- Timestamp – The date and time of the request.
- Department – The hospital department where the request was made.
- Room – The specific room where the task was needed.
- Requirement – The type of request or task.
- Status – The completion status of the task.

1.5.3 Data Volume

The dataset included several thousand rows (exact number to be filled in before submission), covering a period of several months up to one year.

1.5.4 Initial Data Processing

To prepare the dataset for analysis, several preprocessing steps were applied:

- Datetime Parsing:
  Raw timestamps (e.g., "2025-03-15 14:22") were converted into standard datetime format to extract relevant components like the hour of the day and day of the week.
- Feature Engineering:
  New columns were created to capture temporal behavior, allowing for analysis of trends by hour and day.
- Patient Classification (Inpatient vs Outpatient):
  Departments were categorized as inpatient or outpatient to better understand workload origins and to differentiate the demand profiles between hospitalization and ambulatory care.
  This classification played a crucial role in uncovering significant differences in request patterns between department types.

2. Data Analysis and Key Findings

2.1 Workload Distribution by Hour and Department Type

To better understand operational load patterns, we analyzed the volume of requests (tasks) by hour, comparing inpatient departments ("admitted patients") with outpatient departments ("ambulatory patients").

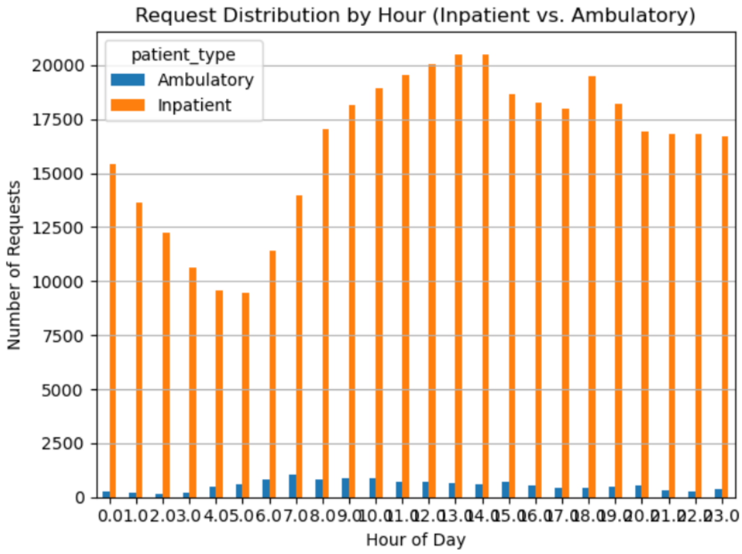
Methodology:

- The dataset was grouped by hour of the day and department type (inpatient/outpatient).
- We calculated the total number of requests for each group.

- A comparative bar chart was created to visualize the differences in hourly workload between the two department types.

Key Insights:

- Inpatient departments dominate the workload across all hours of the day. These departments consistently generate a significantly higher number of requests than outpatient departments.
- Peak hours for inpatient departments are between 09:00 and 16:00, corresponding to intensive medical activity such as rounds, tests, and shift changes.
- Unexpectedly high loads were also observed during late-night hours (00:00–05:00), likely due to urgent testing, cleaning, and patient transport tasks.
- Outpatient departments displayed a relatively low and stable volume of requests, with no significant spikes throughout the day.



Operational Implications:

- Workforce planning should focus primarily on inpatient departments, as they represent the bulk of hospital operational load.
- Staff reinforcements are recommended during peak hours (09:00–16:00), and night shifts should be optimized to accommodate consistent overnight activity.
- Outpatient departments may not require significant staffing adjustments, aside from specific exceptions.

2.2 Workload Distribution by Day of the Week

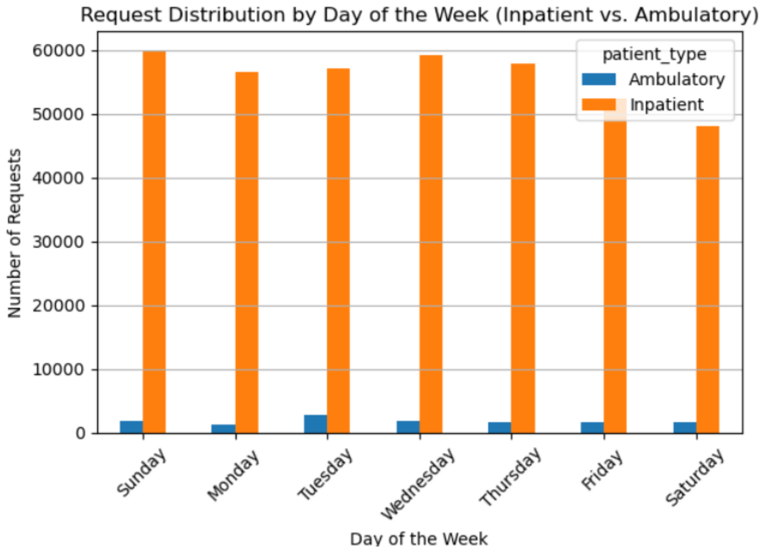
To assess weekly operational trends, we analyzed the number of service requests submitted each day, comparing inpatient and outpatient departments.

Methodology:

- Requests were grouped by day of the week and department type (inpatient vs. ambulatory).
- The total number of requests was aggregated per day for each group.
- A bar chart was created to visualize daily workload patterns.

Key Insights:

- Inpatient departments consistently generate higher workload throughout the week compared to outpatient units.
- Sunday (the first workday in Israel) shows the highest volume of requests in inpatient departments, likely due to accumulated needs over the weekend.
- Other high-load days include Tuesday, Wednesday, and Thursday, indicating mid-week operational intensity.
- Although workload slightly decreases on Fridays and Saturdays, the request volume remains substantial, highlighting the need for continuous staffing even on weekends.
- Outpatient departments show a significantly lower and stable request rate across the week.



Operational Implications:

- Staffing plans should prioritize inpatient departments, especially on Sundays and mid-week (Tuesday to Thursday).
- Weekend coverage must not be overlooked, as operational demands persist even during non-standard workdays.
- Outpatient departments may require minimal adjustments, but periodic checks should be conducted to address any exceptions or special cases.

2.3 Heatmap Analysis: Requests by Day and Hour

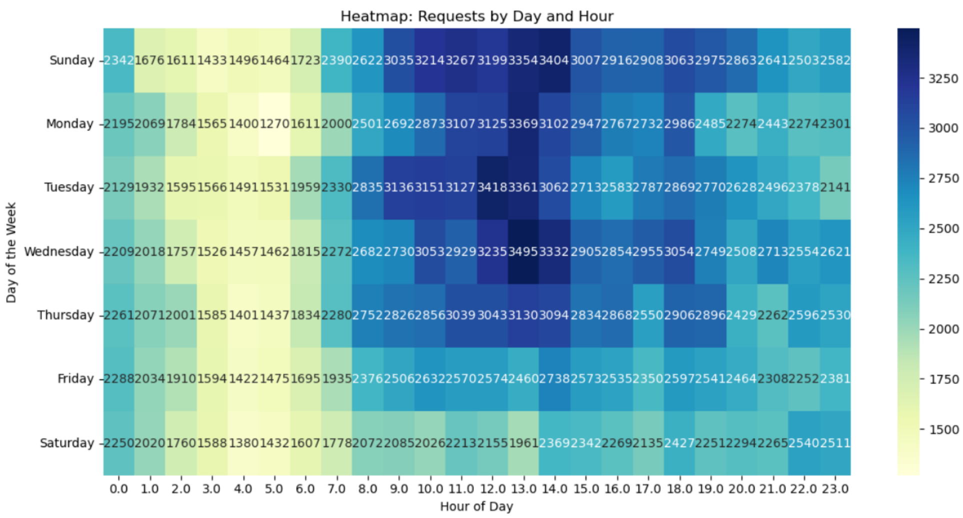
To uncover more granular workload patterns, we analyzed the volume of requests across each hour of the day and each day of the week. This two-dimensional view reveals temporal trends that support dynamic resource allocation.

Methodology:

- Requests were grouped by day of the week (Sunday–Saturday) and hour of the day (0–23).
- A heatmap was generated, where:
  - Darker colors indicate higher request volumes.
  - Lighter colors indicate lower volumes.
  - Numerical values on each cell represent the exact number of requests.

Key Insights:

- Peak activity consistently occurs between 08:00 and 15:00, regardless of the day, with some hours exceeding 3,000 requests.
- Unexpected nighttime activity was observed between 00:00 and 02:00, suggesting the need for overnight staff planning even during presumed "quiet" hours.
- Sundays show high demand from early morning through late afternoon, likely due to weekend backlog.
- Tuesdays, Wednesdays, and Thursdays also show strong morning-to-afternoon activity.
- Evening hours (17:00 and later) reflect a gradual decline in activity, especially on Fridays and Saturdays.



Operational Implications:

- Reinforce staff coverage during peak hours (08:00–15:00) with a focus on high-demand services such as blood tests, cleaning, and consultations.
- Allocate additional staff between 00:00 and 02:00, particularly in inpatient units.
- Prepare for elevated demand on Sundays and midweek, while allowing reduced staffing in the evening and late-night hours on weekends.

2.4 Request Status Distribution by Patient Type

Understanding the distribution of request statuses helps evaluate how tasks are managed operationally in different department types.

Methodology:

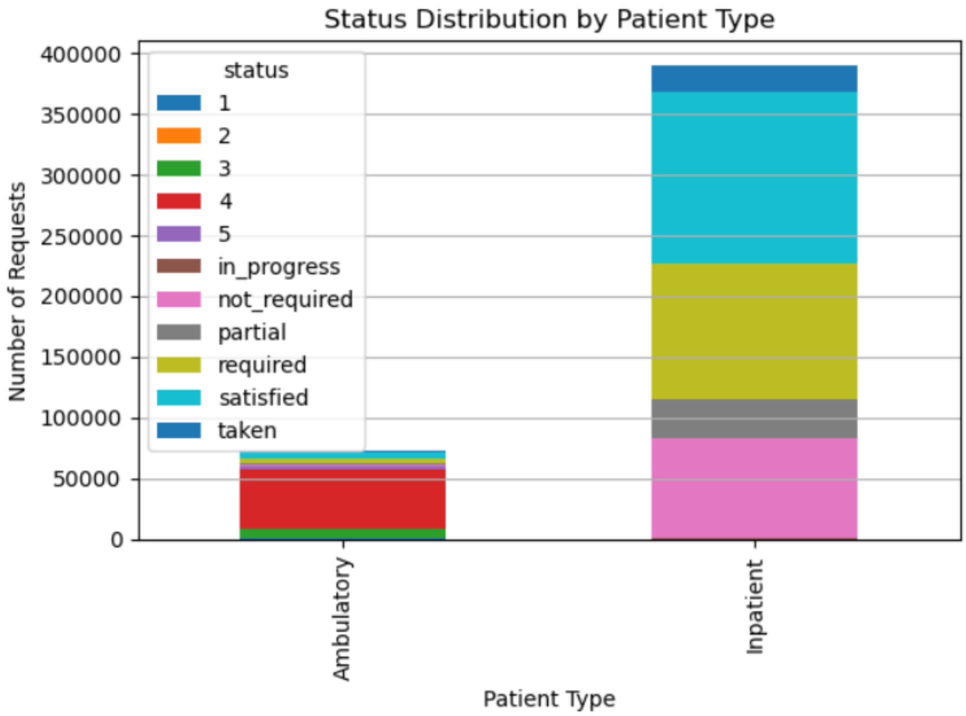
- All service requests were grouped by patient type (Inpatient vs. Ambulatory).
- Status labels (e.g., taken, required, satisfied, partial, not_required) were aggregated and visualized using a stacked bar chart.

Key Insights:

- Inpatient departments exhibit a wide range of statuses, indicating dynamic and ongoing task management.

- High volumes of requests are marked as taken, required, and satisfied – reflecting active task execution.
- A notable number of requests remain in partial or pending states, suggesting the need for more efficient task closure.

- Ambulatory departments, on the other hand, show a dominant proportion of not_required statuses, with relatively few requests needing intervention.



Operational Implications:

- Inpatient units require continuous monitoring and real-time workflow management to avoid bottlenecks and improve task completion.
- The large volume of pending or partial statuses indicates an opportunity for process automation or smarter alert systems.
- Ambulatory departments may benefit from streamlining request entry, as many requests appear unnecessary or redundant.

2.5 Top Request Types by Department

To identify the main drivers of operational load, we analyzed the five most requested task types in the five busiest hospital departments.

Methodology:

- Request volumes were aggregated for each department.
- The top five departments with the highest total number of requests were identified.
- For each of these departments, the five most frequent request types were extracted and ranked.

Key Insights:

- Department A has the highest overall request volume. Its most common requests are:
  - Blood Tests – 52,212 requests

  - Doctor Visit – 30,135 requests
  - Cleaning – 27,833 requests
  - Bed Occupancy – 25,976 requests
  - ECG Test – 19,936 requests
- Department B also shows high activity:
  - Blood Tests – 35,040
  - Cleaning – 18,938
  - Doctor Visit – 16,654
  - Bed Occupancy – 13,087
  - ECG Test – 4,810
- Department C has similar patterns:
  - Doctor Visit – 26,873
  - Blood Tests – 26,130
  - Cleaning – 20,833
  - Bed Occupancy – 17,968
  - ECG Test – 8,923
- Trauma Unit (ER Division A) and Ambulatory ER both show high volumes of "Urgency Level" tasks:
  - ER Division A – 10,218 urgency-level tasks
  - Ambulatory ER – 26,734 urgency-level tasks

Operational Implications:

- Blood tests, doctor visits, and cleaning tasks represent a significant share of all requests. These recurring services should be assigned dedicated teams, with tailored training and optimized schedules.
- ER departments (both inpatient and ambulatory) rely heavily on triage and urgency-related requests. These should be supported by specialized staff and real-time response protocols.
- Department A stands out with the heaviest workload, especially in lab and physician-related requests. It may require enhanced infrastructure, staff reinforcement, and targeted operational improvements.

2.6 Correlation Analysis

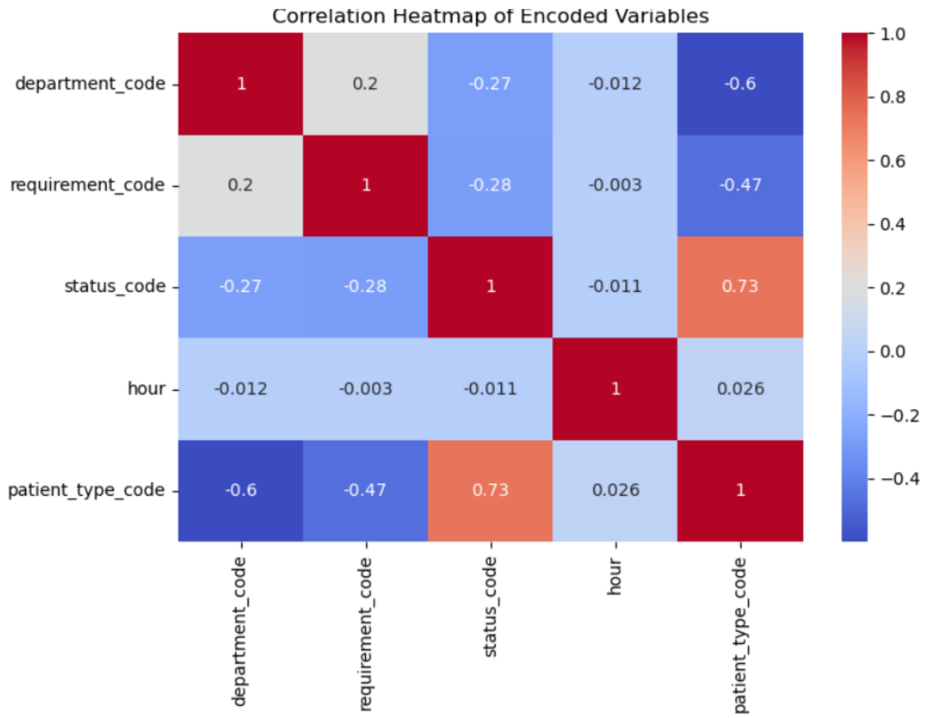
To uncover relationships between key variables in the dataset, we conducted a correlation analysis across patient type, department, request type, request status, and time-related attributes.

Methodology:

- Categorical variables (e.g., department, request type, status, patient type) were encoded numerically.
- A correlation matrix was computed using Pearson's method to quantify the strength and direction of linear relationships between variables.

Key Insights:

- Strong positive correlation (0.72) between patient type and request status:
  Inpatient requests are more likely to be in active statuses such as taken, required, or partial, whereas ambulatory requests are mostly not_required.
- Moderate negative correlation (-0.59) between patient type and department:
  Certain departments are strongly associated with one patient type (e.g., Department A is mainly inpatient).
- Moderate-to-weak negative correlation (-0.45) between request type and patient type:
  Some task types are more typical for inpatients (e.g., blood tests), while others are more common in ambulatory care.
- Weak correlations (~ -0.25) between request status and department or request type:
  Statuses vary slightly by department and request, but no strong pattern is found.
- No significant correlation between hour of day and other variables:
  While workload volume fluctuates by hour, the request type or status does not directly depend on the time.



Operational Implications:

- Patient type is a key driver of task status and behavior. Workflow design should differentiate between inpatient and ambulatory services to reflect their distinct operational characteristics.
- Department-specific strategies are necessary, as departments are aligned with particular request types and load patterns.
- Tailored workflows and staff assignments based on request type can help reduce unnecessary or incomplete tasks, particularly for common inpatient requests.
- Temporal optimization should rely on volume analysis (e.g., heatmaps), rather than on correlation with status or type, due to the weak direct relationship.

2.7 Heatmap Analysis: Request Type by Hour

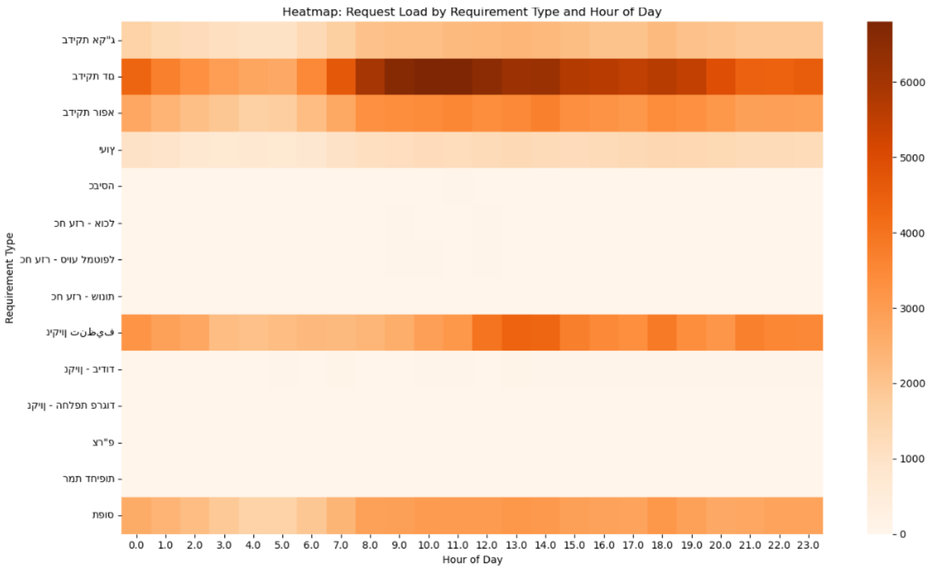
To optimize hourly staff allocation and operational readiness, we analyzed the distribution of requests by type across all 24 hours of the day.

Methodology:

- The six most frequent request types were selected based on total volume.
- Requests were grouped by hour of the day (0–23) and type of service.
- A heatmap was generated showing the number of requests per task per hour.

Key Insights:

- Urgency Level requests dominate all hours, with peaks between 08:00 and 20:00, indicating continuous triage-related activity.
- ECG Tests are evenly distributed, with moderate peaks around 08:00–17:00.
- Bed Occupancy requests are concentrated in late-night (22:00–03:00) and early-morning hours, reflecting off-hour admissions and discharges.
- Consultations display moderate and consistent load, with a peak around 12:00–15:00.
- Cleaning tasks show two main peaks: early morning (06:00–09:00) and evening (18:00–21:00), corresponding to standard cleaning routines.



Operational Implications:

- Staff assignments should align with the hourly peaks of specific request types. For instance:
  - Morning shifts should prioritize cleaning and blood tests.
  - Midday staffing should focus on consultations and ECG tests.
  - Night shifts must be prepared to handle bed management and urgent care.
- Resources can be allocated more efficiently by using real-time workload forecasting based on these patterns.

2.8 Heatmap Analysis: Request Type by Day of the Week

To enhance weekly workforce planning, we analyzed the distribution of request types across days of the week, identifying which tasks peak on specific days.

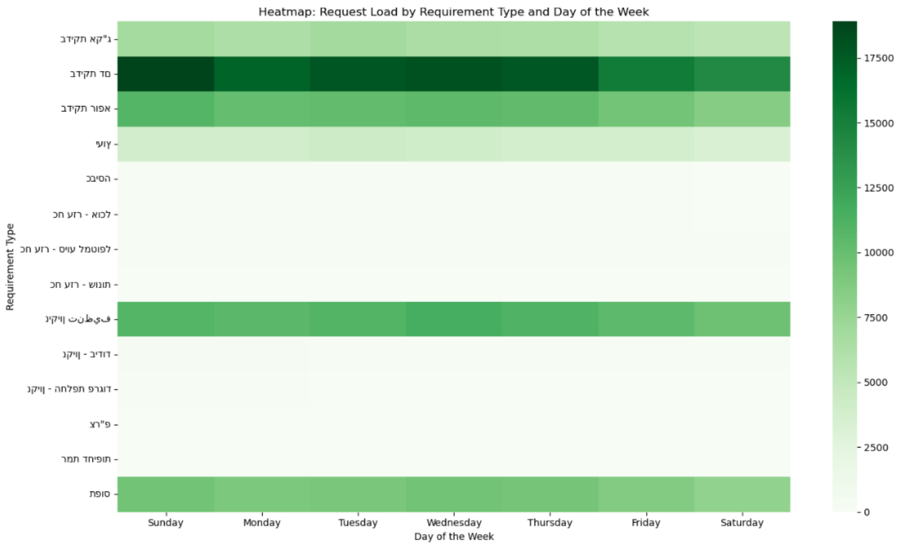
Methodology:

- The most common request types were selected.
- Request volumes were grouped by day of the week (Sunday–Saturday) and request type.
- A heatmap was generated to highlight day-specific demand patterns.

Key Insights:

- Urgency Level tasks remain consistently high throughout the week, with Sunday being the most demanding day.
- ECG Tests show stable volume across all weekdays with no significant fluctuations.

- Bed Occupancy requests increase slightly towards weekends (Friday–Saturday), potentially due to increased patient turnover.
- Cleaning tasks peak slightly on Sunday and Monday, likely as part of post-weekend routines.
- Consultation loads remain stable, showing no sharp variations throughout the week.



Operational Implications:

- Sunday requires reinforced staffing, particularly for urgent care tasks.
- Weekend planning should consider increased needs for bed management and cleaning.
- Stable tasks like consultations and ECG tests can be handled with consistent scheduling.
- Weekly shift planning should remain flexible, adjusting for predictable demand variations across task types.

2.9 Smart Staff Scheduling Recommendations

To support data-driven workforce planning, we compiled targeted staffing recommendations based on request patterns segmented by patient type, request type, day of the week, and hour of the day.

Methodology:

- The dataset was grouped by patient type (inpatient or ambulatory), request type, weekday, and hour.
- For each unique combination, we calculated the exact number of requests.
- A summary table was generated to highlight the peak demand time for each request type and patient group.

Key Insights:

- Each request type has a distinct load pattern, varying significantly by both day and time.
  - Example: Blood test requests for inpatients peaked on Sunday at 09:00, while cleaning tasks peaked during late evening hours.
- Inpatient departments showed higher overall volume, with multiple peak periods per request type.
- Ambulatory departments had fewer high-load periods, but still exhibited specific timing patterns for certain requests.

	Patient Type	Requirement Type	Day	Hour	Request Count
0	Inpatient	בדיקת דם	Sunday	11	1101
1	Inpatient	בדיקת דם	Sunday	10	1089
2	Inpatient	בדיקת דם	Tuesday	10	1086
3	Inpatient	ניקיון تنظيف	Wednesday	13	720
4	Inpatient	ניקיון تنظيف	Tuesday	13	664
5	Inpatient	ניקיון تنظيف	Monday	13	637
6	Inpatient	בדיקת רופא	Sunday	14	598
7	Inpatient	בדיקת רופא	Sunday	9	591
8	Inpatient	בדיקת רופא	Monday	14	570
9	Inpatient	תפוס	Wednesday	13	522
10	Inpatient	תפוס	Tuesday	13	506
11	Inpatient	תפוס	Sunday	14	503
12	Inpatient	בדיקת אק"ג	Sunday	13	387
13	Inpatient	בדיקת אק"ג	Monday	13	385
14	Inpatient	בדיקת אק"ג	Sunday	14	382
15	Inpatient	יעוץ	Sunday	20	244
16	Inpatient	יעוץ	Wednesday	18	239

Operational Implications:

- Create time-specific staff schedules based on the actual peak load hours for each department and task type.
- Establish reinforcement teams that activate automatically during forecasted peak times (e.g., Monday mornings for blood tests).
- Move away from fixed shifts toward dynamic shift planning that adapts to real-time needs.
- These insights enable proactive resource allocation, reducing delays and bottlenecks.

2.10 Request Load by Hour and Request Type, Split by Patient Type

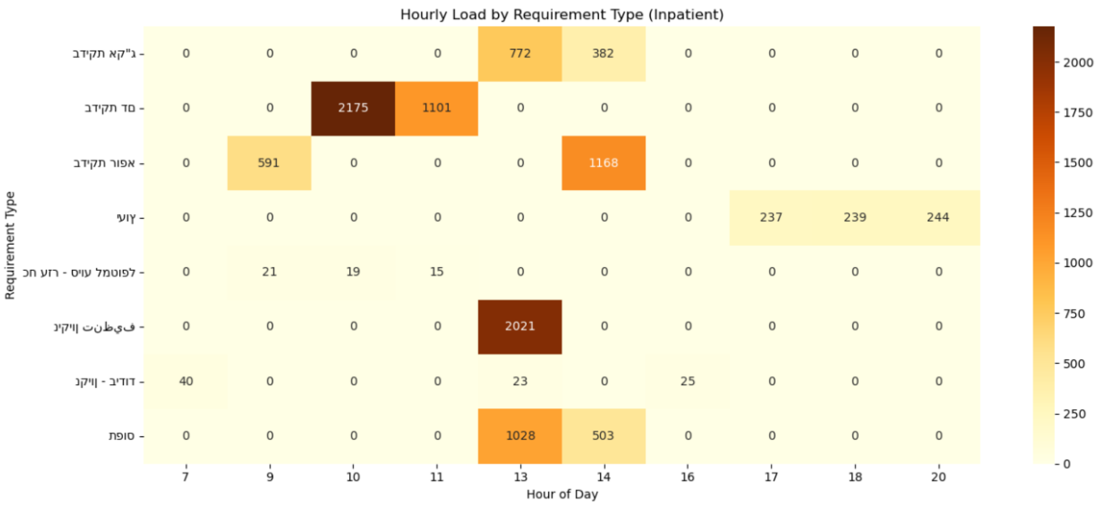
To gain deeper operational insights, we analyzed the hourly volume of each request type separately for inpatient and ambulatory departments. This comparison helps distinguish the unique patterns and needs of each patient group.

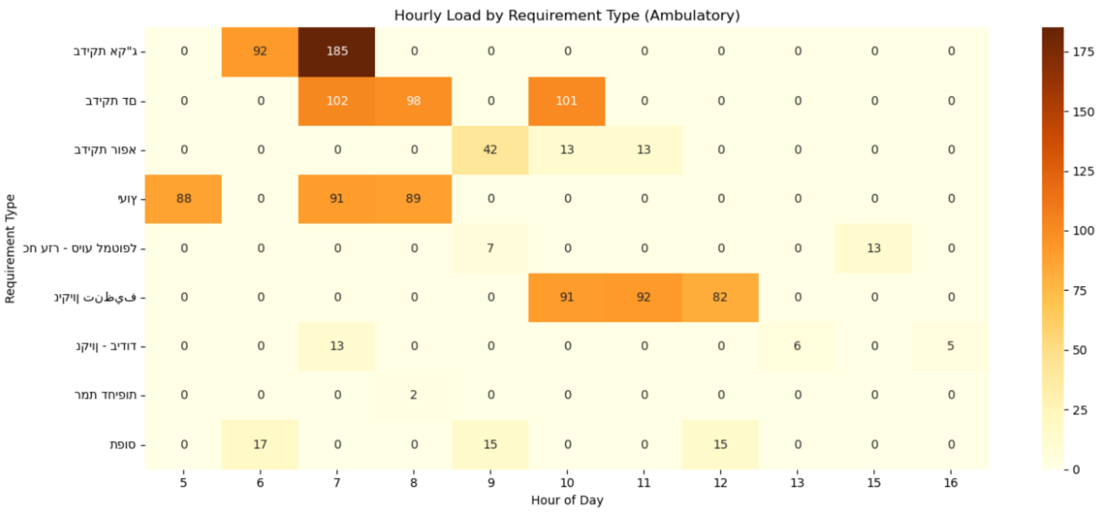
Methodology:

- The top request types were selected based on overall frequency.
- For each request type, we calculated hourly request volumes separately for inpatients and ambulatory patients.
- The results were visualized in side-by-side bar charts for each patient type.

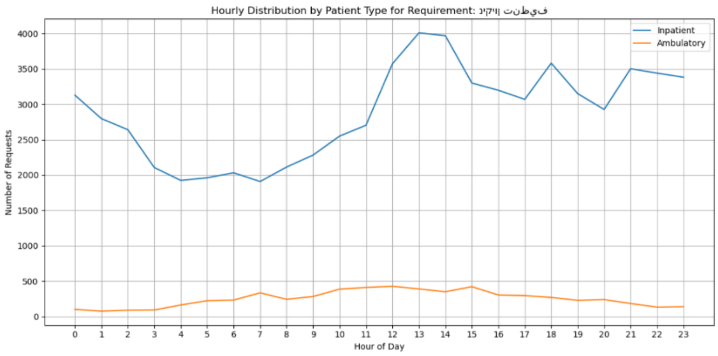
Key Insights:

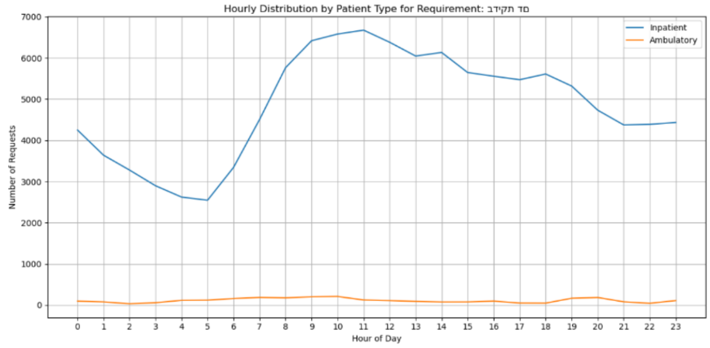
- Inpatient departments consistently show higher request volumes across all hours and all task types.
- Their peak hours vary by task: e.g., blood tests peak at 09:00, while cleaning peaks in early evening.
- Ambulatory departments have a narrower request window, concentrated during standard daytime hours (08:00–16:00).
- Tasks like consultations and diagnostics dominate in ambulatory care, with limited overnight activity.
- Some tasks (e.g., urgency-level assessments) appear in both department types but are significantly more frequent in inpatient settings.

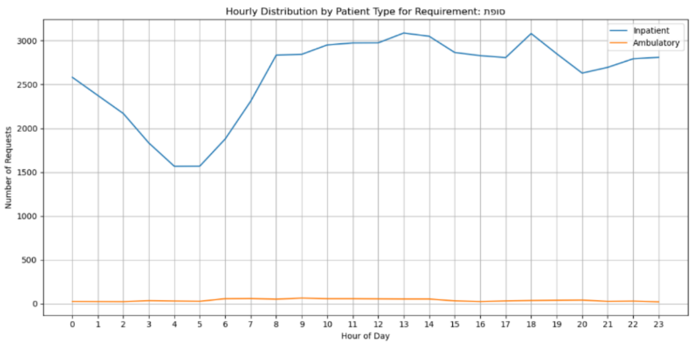


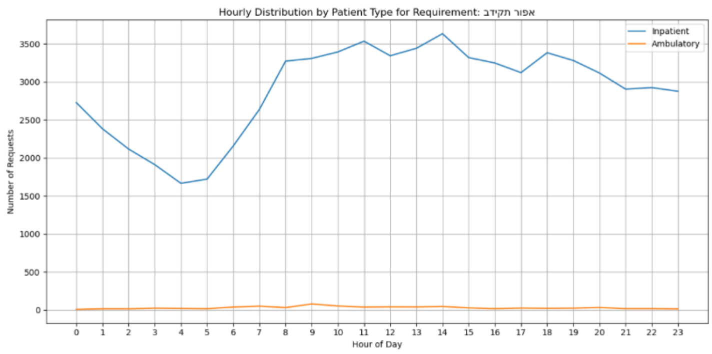


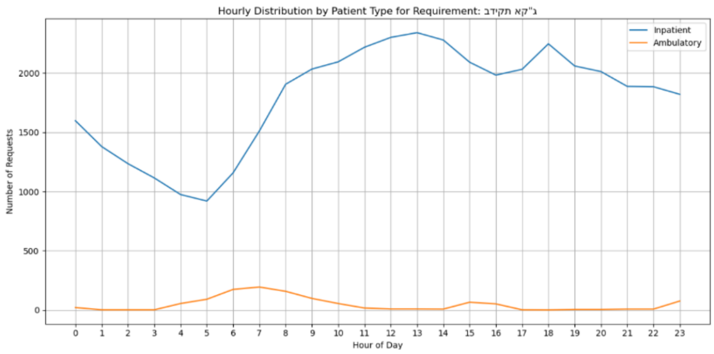
Hourly Workload Comparison Between Inpatient and Ambulatory Patients for Top Request Types:











2.10.1 Hourly Trends by Request Type

This subsection presents hourly trends for each of the most frequent request types (e.g., blood tests, cleaning, ECG).
Line charts were used to visualize how the volume of requests fluctuates by hour and differs between inpatient and ambulatory departments.

These visualizations reveal the unique temporal patterns of each task, supporting more accurate, task-specific scheduling decisions.

2.10.2 Focused Comparison: Inpatient vs. Ambulatory for Key Tasks

This section compares hourly request volumes for two high-demand services - “Order Arrangement” and “Urgency Level” - across inpatient and ambulatory departments.
The line charts demonstrate consistent differences in demand intensity and timing between the two patient types, highlighting the need for patient-specific scheduling approaches.

2.10.3 Operational Patterns by Task Type

To support precision staffing, we analyzed hourly workload patterns for top request types by patient type. The table below summarizes the key trends and operational implications observed from the line charts:

Request Type	**Inpatient Trend**	**Ambulatory Trend**	**Operational Insight**
Blood Test	Significant morning load (09:00–12:00), then gradual decline	Low but steady volume, with mild peak at 07:00–10:00	Reinforce morning staff in inpatient units
Physiotherapy / Viking	Moderate and stable load from 10:00–15:00	Consistent levels throughout the day	Focus staffing between 11:00–14:00
“Proper Makeup”	Peak between 09:00–14:00, then decline	Very low number of requests	Prioritize early handling in inpatient units
“Soft”	Peaks between 10:00–15:00	Balanced, low-to-moderate activity throughout the day	Balance shifts around late morning and early afternoon
Medical Check	High morning-to-noon load	Sharp peak between 06:00–08:00, then drop	Early morning support for ambulatory units
Consultation	Gradual increase throughout the day, peaking in afternoon	Light load, mainly between 06:00–08:00	Prepare for afternoon load in inpatient departments
“Doviz-Viking”	Very low and inconsistent	Few, irregular requests	No special staffing needed
Eye Screening	Irregular sharp peaks between 09:00–13:00	Stable throughout the day	Targeted readiness for inpatient peak periods
General Treatment	Low, scattered activity throughout the day	Same	No reinforcement required

Operational Implications:

- Scheduling and task management must account for different temporal patterns between inpatient and ambulatory services.
- Night shifts are more critical for inpatient units, while ambulatory departments may function efficiently with standard daytime staffing.
- By splitting the analysis by patient type, staff can be assigned more accurately, minimizing idle time and overloads.

2.11 Summary of Operational Bottlenecks

Based on the comprehensive analysis, we identified the key operational bottlenecks that impact hospital efficiency:

1. Temporal Load Peaks

- Peak activity occurs between 08:00 and 15:00 across nearly all request types and days.

- Unexpected load spikes during late-night hours (00:00–02:00) – particularly in inpatient departments – suggest the need for better night-shift readiness.

2. Inpatient vs. Ambulatory Gaps

- Inpatient units handle a significantly higher and more diverse volume of requests.
- Ambulatory departments operate mostly during standard business hours, allowing for more predictable scheduling.

3. Task-Specific Patterns

- Certain services (e.g., blood tests, cleaning, and bed occupancy) exhibit clear hourly and daily demand patterns, requiring tailored shift planning.
- Others (e.g., consultations, ECGs) show more stable demand, suitable for routine scheduling.

4. Status Disparities

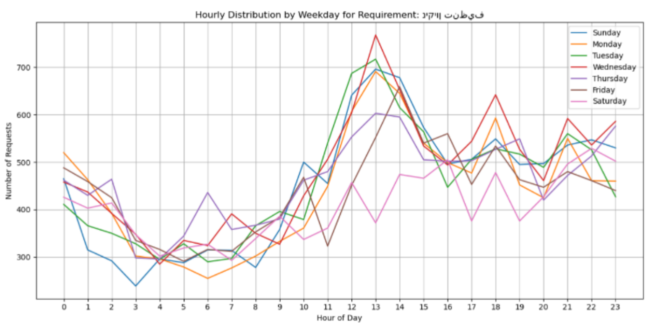
- High presence of incomplete or partially handled requests in inpatient departments indicates potential inefficiencies in follow-up or task closure.

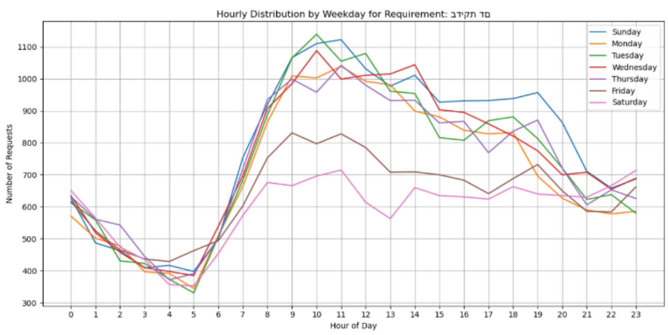
5. Staffing Misalignment

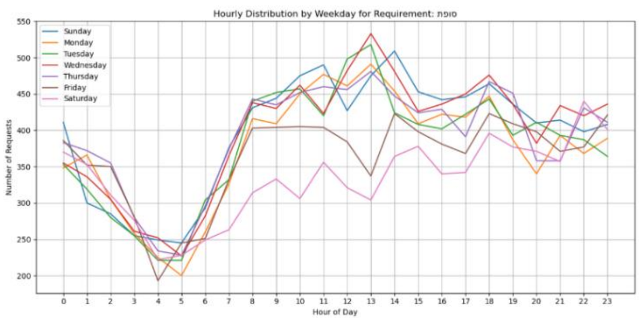
- Current workforce planning may not align with true workload distribution, especially during off-hours and for high-volume request types.

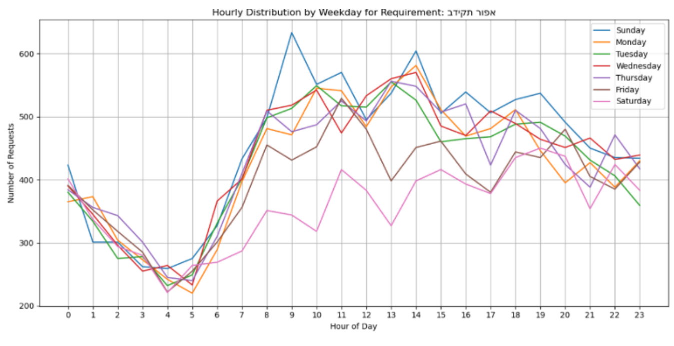
Strategic Recommendations:

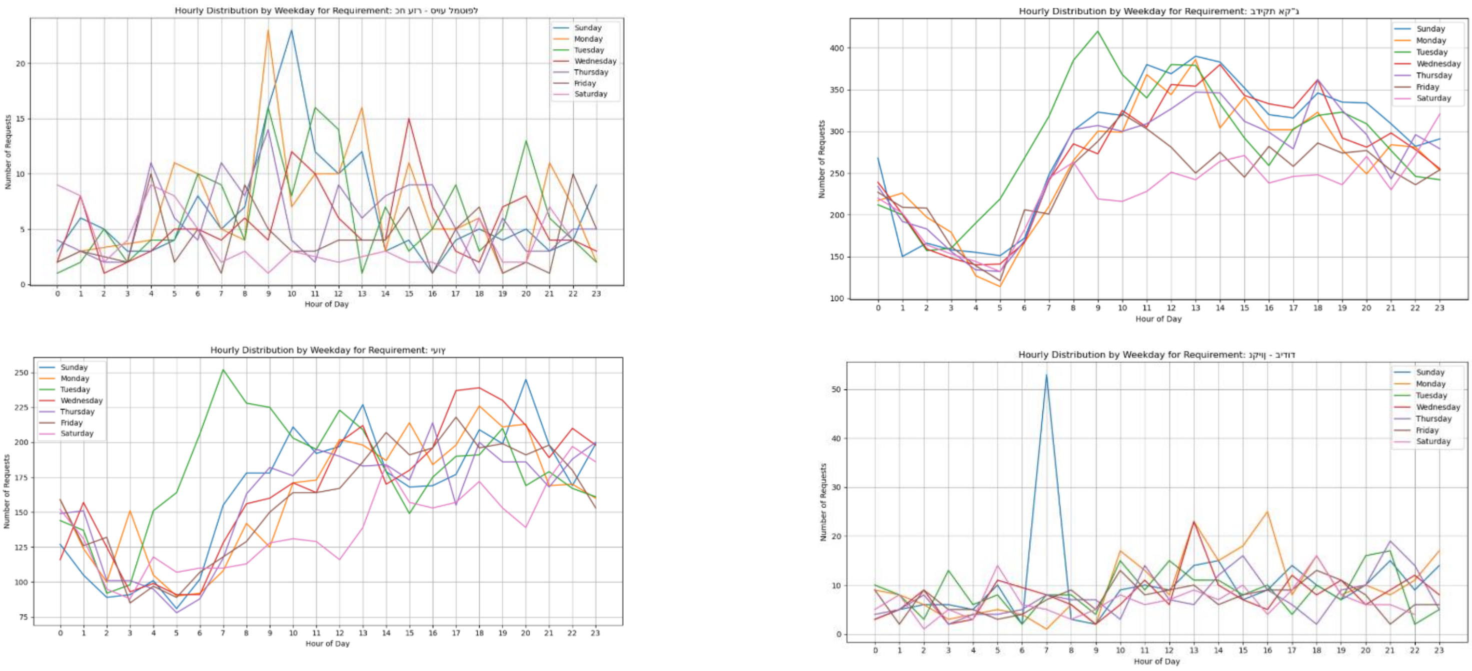
- Move from static shift models to data-driven, dynamic scheduling.
- Implement real-time monitoring and alerts for high-load periods.
- Prioritize cross-functional staff capable of handling multiple task types during peak times.











Hourly request volume trends for top task types, segmented by day of the week. The patterns highlight critical load windows and support precise bottleneck identification.

2.12 Success Rate by Request Type and Department

Objective:
This analysis aims to measure the success rate for each combination of request type and department. A request is considered "successful" if its final status is marked as "satisfied".

Understanding success rates helps reveal which departments and request types are effectively fulfilled and which ones may require operational improvements.

Findings:
High success rates were observed in the following combinations:

- Isolation Cleaning Requests in the Ambulatory ER department – success rate of 60.3%.
- Privacy Curtain Replacement and other cleaning-related requests in the same department – over 55%.
- Departments A and Trauma New also showed strong performance (~54%) in certain tasks.

Interpretation:
These tasks are likely well-defined, logistically supported, and easy to prioritize, which results in high completion rates.

Low success rates were observed in:

- Blood tests in Departments A and C – only ~17% success.
- Eye screenings in the Ambulatory department – just 9.6%.
- Doctor Visits in the New Triage ER – 0% success rate.

Interpretation:
These may reflect unfulfilled or delayed tasks, registration issues, or a lack of responsibility clarity. Investigating workload, processes, and staffing in these areas is recommended.

```
Success Rate by Requirement and Department:
status                                  satisfied
requirement           department
0.602941  מלר"ד אגף מהלכים        נקיון - בידוד
0.560000  מלר"ד אגף מהלכים מלר"ד פרגוד החלפת - נקיון
                      A                   0.558824
                      0.548387     טריאז' מיון חדש
0.545455        טראומה חדש        נקיון - בידוד
...                                            ...
בדיקת דם             A                  0.172278
                      C                   0.172254
0.095771  מלר"ד אגף מהלכים                 יעוץ
0.000000   טריאז' מיון חדש          בדיקת רופא
                      0.000000              דימות

[74 rows x 1 columns]
```

Key Takeaways:

- High-performing departments and tasks can serve as models for replication in other areas.
- Low-performing combinations represent bottlenecks or weak points in hospital workflows.
- These insights can inform resource allocation, staff reinforcement, and process redesign, especially in the ER and high-volume departments.

2.13 Identification of Departments with High Cancellation or Incompletion Rates

Objective:
This analysis focuses on identifying hospital departments with a high percentage of service requests that were either not completed, canceled, or left in intermediate statuses (such as partial, not required, or in progress).
The goal is to uncover potential operational gaps or workflow inefficiencies.

Findings:

Several departments demonstrated significant rates of uncompleted or canceled requests:

- Department C and Department A exhibited a high share of blood test and medical consultation requests left in partial or not required status.
- In the New Triage ER, a considerable portion of doctor visit requests were never marked as completed, raising concerns about process execution or tracking.
- The Ambulatory ER showed a large volume of “not required” statuses, particularly for diagnostic-related tasks.

```
Departments with High 'not_required' or 'partial' Counts:
status              not_required  partial
department
A                        36298.0  14375.0
B                        23292.0   9728.0
C                        21405.0   7352.0
399.0      1818.0        מלר"ד אגף מהלכים
231.0      307.0               טראומה חדש
0.0        5.0            טריאז' מיון חדש
0.0        4.0                      דימות
```

Interpretation:

- High incompletion or cancellation rates may result from:
  - Task abandonment due to low prioritization
  - Systemic delays or overloads
  - Lack of task ownership or handoff clarity
- These trends might reflect areas where tasks are logged but not actively managed, or where processes are not fully integrated with staff routines.

Recommendations:

- Conduct targeted process audits in the identified departments to evaluate where requests are dropped.
- Consider adding alerts or escalation mechanisms for tasks left unresolved beyond a certain time threshold.
- Improve accountability by linking tasks to specific roles or staff members responsible for completion.

2.14 Room-Level Analysis: Top Rooms by Number of Requests

Objective:
The purpose of this analysis is to identify rooms that consistently receive a high number of service requests over time. This insight helps locate operational pressure points within departments and may reveal patient clusters with high care needs (e.g., complex or critical cases), or potential issues in room-related logistics.

Key Findings:

The table below presents the ten rooms with the highest number of requests:

```
Top Rooms by Number of Requests:
room
A103    11029
A104    11000
A114    10779
A115    10775
A113    10434
A105    10364
A112    10265
A116     9957
A110     9751
A102     9682
Name: count, dtype: int64
```

These rooms are all located in Department A, indicating that this department experiences consistently high demand. The top three rooms (A103, A104, A114) show particularly elevated volumes.

Interpretation:

- Room A103 stands out with over 11,000 requests, which may point to:
  - High patient turnover
  - Placement of complex patients (e.g., intubated, elderly, severely ill)
  - Or the need to review room-specific logistics or infrastructure

- Conversely, room A102 shows a lower request count, possibly indicating fewer high-needs patients or more efficient operations.

2.15 Request Volume Analysis: Weekdays vs. Weekends

Objective:

The goal of this analysis was to compare the volume of service requests by request type between weekdays (Sunday–Thursday) and weekends (Friday–Saturday).
This distinction helps identify which tasks require continuous support throughout the week, versus those that occur exclusively on weekdays.

Key Findings:

The request types with the highest weekend activity were:

- Blood Tests – over 87,000 requests on weekends alone.
- Cleaning / General Cleaning – more than 52,000 weekend requests.
- Doctor Visits, Room Occupancy (Taken), and ECG Tests – tens of thousands of requests each, even during weekends.

Conversely, certain tasks showed very low activity on weekends:

- Patient Assistance, Curtain Replacement, Laundry, Support Staff – Food, and X-ray Transport – each had fewer than 500 requests on weekends.

```
Request Volume: Weekday vs Weekend:
                         Weekday  Weekend
requirement
28598     87372                  בדיקת דם
19488     52611            ניקיון تنظيف
18063     52041                בדיקת רופא
16119     45321                      תפוס
10942     32111               בדיקת אק"ג
6862      19469                      יעוץ
322       1071              נקיון - בידוד
173       584     כח עזר - סיוע למטופל
153       403                     כביסה
112       440              כח עזר - אוכל
80        290      נקיון - החלפת פרגוד
12        69              כח עזר - שונות
7         33                       צר"פ
```

Conclusions and Operational Insights:

Critical medical and operational services (e.g., blood tests, cleaning, ECG) remain highly active throughout the week and must be staffed continuously, including on Fridays and Saturdays.

Supportive and service-oriented tasks (e.g., food delivery, laundry, personal assistance) are concentrated on weekdays, with minimal weekend presence.

Resource Allocation Insight:
Shift planning can be optimized by ensuring core medical staff are present seven days a week, while support staff scheduling can be concentrated during weekdays to match task demand.

2.16 Analysis of Repeated Requests by Room and Requirement

Objective:

This analysis aimed to identify which room-and-requirement combinations appeared repeatedly, in order to detect recurring patterns of clinical or operational needs.
Such patterns may indicate bottlenecks, systemic inefficiencies, or specific areas that require reinforcement or process review.

Key Findings:

- The most frequently repeated combination was blood test requests in rooms A103, A104, A115, A114, and A105, each with more than 3,000 occurrences.
  This suggests sustained clinical workload in these locations and may reflect long-term care or high-dependency patients.
- At the lower end of the list, less common combinations included:
  - Laundry and patient assistance in room C308
  - Other recurring combinations such as doctor visits, curtain replacement, and ECG appeared 21–22 times.

```
Repeated Requests (Room + Requirement) > 20 Times:
      room           requirement  count
102   A103              3484    בדיקת דם
113   A104              3481    בדיקת דם
237   A115              3402    בדיקת דם
225   A114              3373    בדיקת דם
123   A105              3291    בדיקת דם
..     ...                   ...     ...
544   C308                 22       כביסה
546   C308  22     כח עזר - סיוע למטופל
742   22    נקיון - החלפת פרגוד     דמים
31     418              21    בדיקת רופא
739   21    נקיון - החלפת פרגוד    א.ק.ג

[453 rows x 3 columns]
```

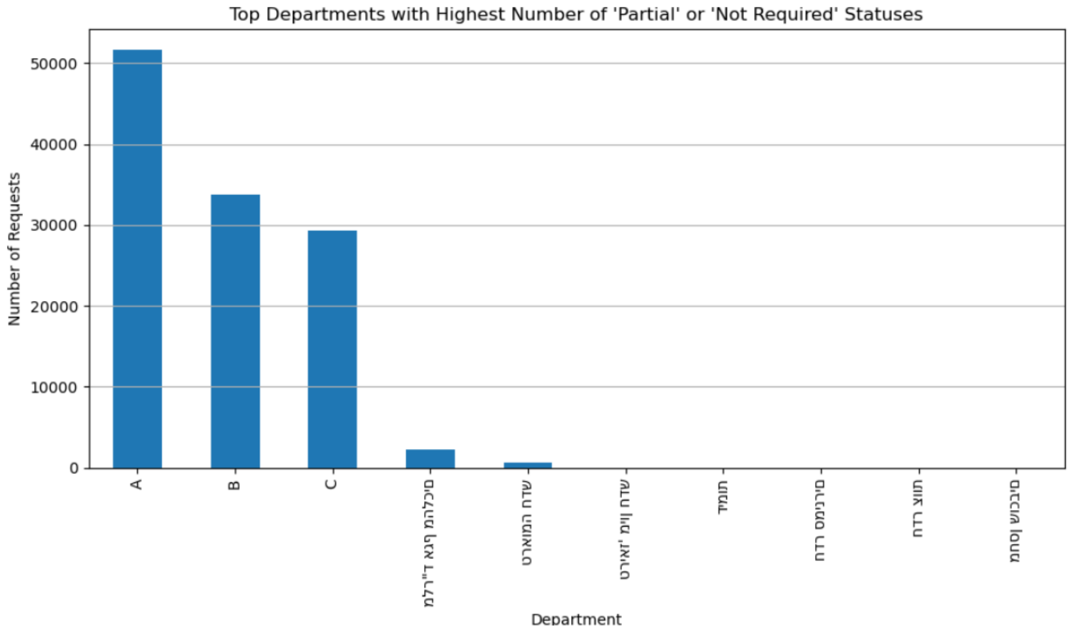
Conclusions:

- Some rooms experience persistently high volumes of repeated clinical requests (especially blood tests), potentially reflecting long-term hospitalization or intensive care usage.
- A repeated concentration of the same requests in specific rooms may indicate operational challenges such as missing equipment, delayed task handling, or inadequate protocols.
- It is recommended to investigate the root causes of repeated requests, particularly when they involve operational tasks like cleaning, patient support, or curtain changes.

2.17 Departments with High Rates of “partial” or “not_required” Statuses

Objective:

This analysis aimed to identify departments with a high number of service requests that ended with the status "partial" or "not_required".
These statuses may indicate incomplete execution, unclear workflows, or redundant requests, and can highlight areas where operational inconsistencies exist.



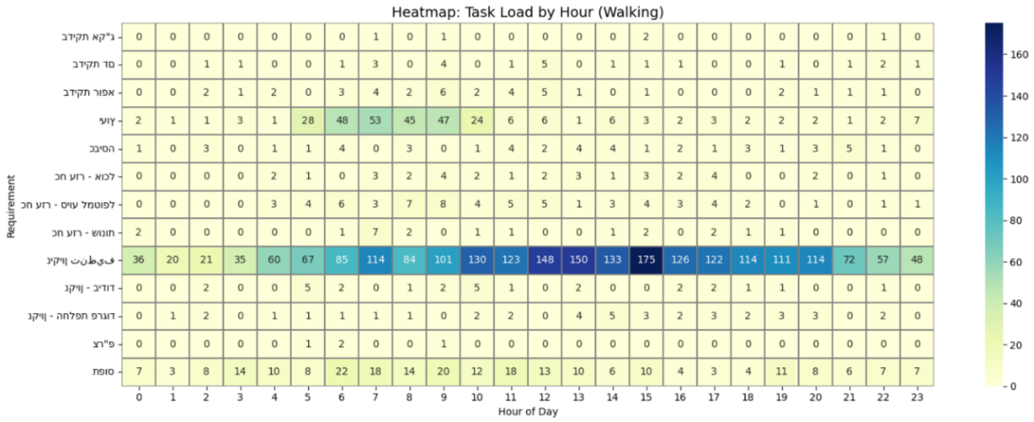
Key Findings:

- Departments A, B, and C had a significantly high number of requests marked as "partial" or "not_required", with tens of thousands of instances in each. This may point to one or more of the following issues:
  - Miscommunication or lack of coordination between teams
  - Redundant or unnecessary requests being submitted
  - Inconsistent execution processes or overlapping task categories
  - Improper documentation of completed tasks in the system
- Other departments such as Ambulatory ER and New Trauma Unit also appeared in the analysis, but with far fewer instances.

Conclusions and Recommendations:

- A thorough review of workflows in Departments A, B, and C is highly recommended:
  - Are there communication gaps between teams?
  - Are submitted requests fully justified?
  - Are some tasks being completed but not properly recorded?
- Consider implementing a training and auditing process to ensure staff understand how to properly open and close requests in the system.
- Combining this analysis with success rate metrics (e.g., satisfied status) can create a more complete picture and guide process improvements.

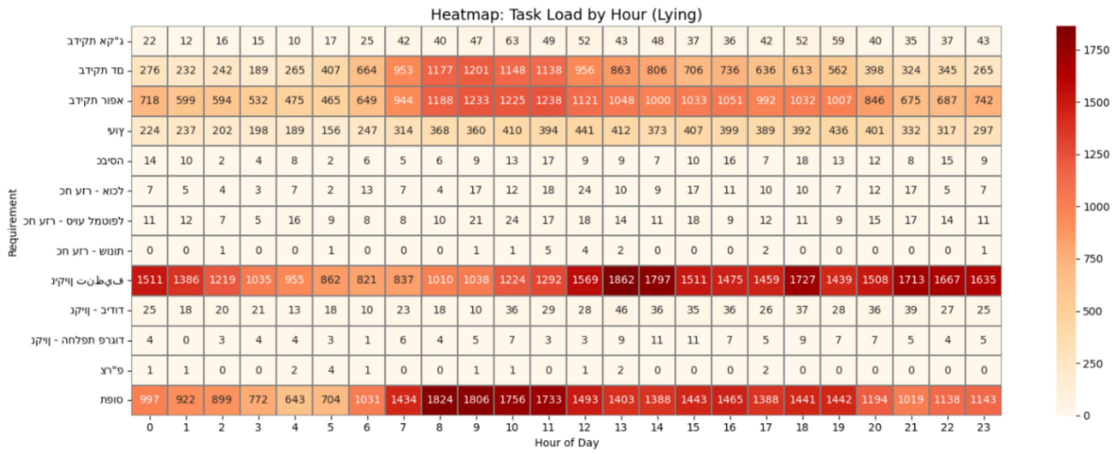
2.18 Heatmap: Task Load by Hour (Walking Patients)



This heatmap shows the volume of requests made by ambulatory patients throughout the day.

- The vast majority of requests fall under vital signs monitoring, with clear peaks between 10:00 and 16:00**.**
- Other request types are relatively infrequent and exhibit low variation.
- This suggests that ambulatory patients generate demand during a narrow range of daytime hours, with minimal activity at night.

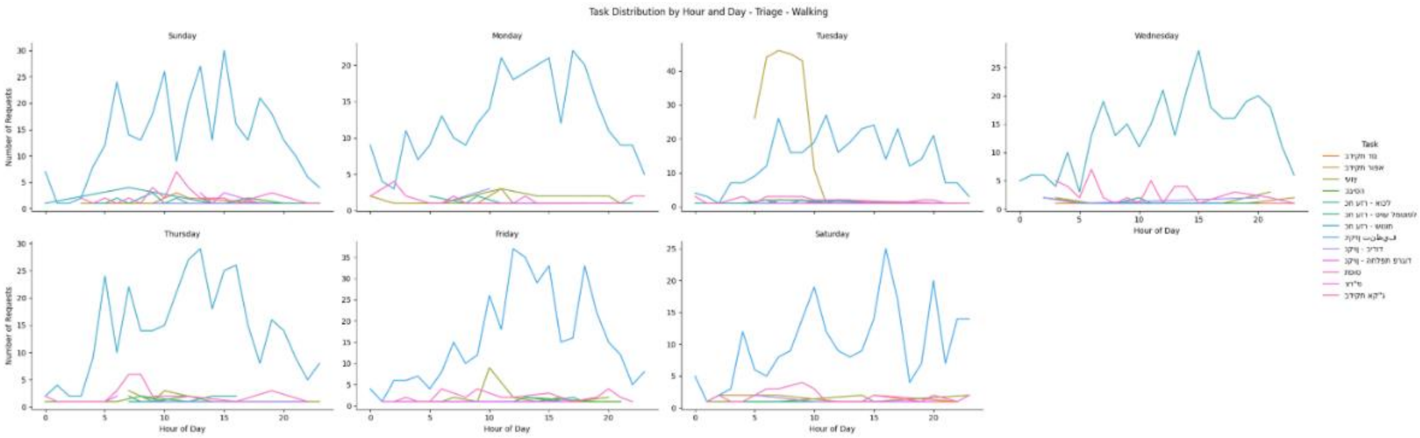
2.19 Heatmap: Task Load by Hour (Inpatients)

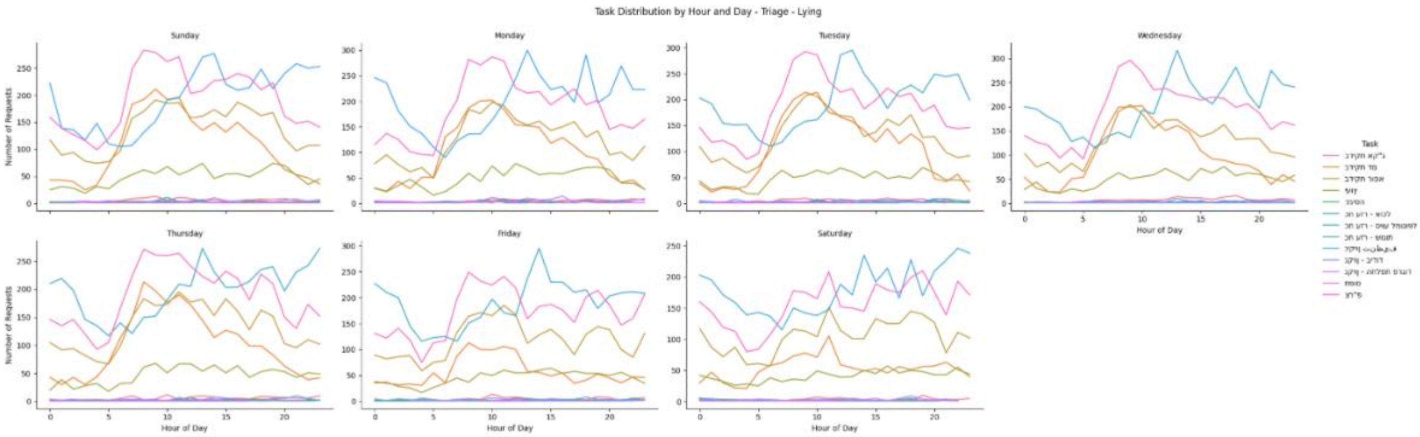


This heatmap displays hourly request volumes from inpatient departments.

- The most common requests are blood tests, physician visits, cleaning, and monitoring.
- The heatmap reveals sustained high demand from early morning (6:00 AM) through late evening (9:00 PM).
- Inpatient care requires consistent staffing across a broader daily window compared to ambulatory services.

2.20 Task Distribution by Hour and Day (Walking vs. Lying Patients)





This set of line charts compares request distribution by hour and weekday, separately for ambulatory and inpatient patients:

- Walking patients: Task patterns fluctuate significantly across the week, with spikes in the mornings.
- Lying patients: Consistent demand throughout the week, especially in essential services.
- The contrast highlights the importance of customized scheduling strategies based on patient type and daily trends.

2.21 Summary- Emergency Department Data Analysis

- The analysis of Emergency Department (ED) requests revealed clear temporal, departmental, and task-related patterns that impact operational performance.
- Time-Based Insights:
  Request volumes peaked between 10:00 and 13:00, especially for ambulatory patients.
  Weekdays (Sunday–Thursday) showed significantly higher volumes than weekends, although weekend demand remained substantial for critical services.
  Heatmaps confirmed that Monday mornings are among the busiest periods.
- Department and Task Patterns:
  Departments A, B, and C emerged as the busiest units, with common request types including blood tests, cleaning, doctor visits, and imaging.
  Repeated requests in specific rooms (e.g., A103–A116) suggested either intensive care routines or recurring service issues.

- Status Analysis:
  Departments A–C had high proportions of "partial" and "not_required" request statuses, indicating potential workflow breakdowns or documentation inconsistencies.
  Success rates varied by task: simple operational tasks were completed more successfully than complex or medical procedures.
- Patient Type Trends:
  Ambulatory patients generated significantly more requests, particularly during peak hours.
  Inpatients submitted requests more evenly across the day but in lower volumes.
- Room-Level Repetition:
  Hundreds of recurring request patterns (same room + same task) were observed, particularly for blood tests, suggesting intensive care routines or repeated task failures.
- Conclusion:
  This analysis highlights the importance of smart workforce planning, task completion monitoring, and targeted quality control in the ED.
  By identifying temporal and spatial bottlenecks, this data-driven approach lays the groundwork for a predictive and adaptive operational model, discussed in the next chapter.

3. Proposed Solution – Forecast-Based Staff Planning in the Emergency Department

3.1 Problem Definition and Goal of the Solution

Hospitals often experience inconsistent and unpredictable workload patterns across departments. Traditional staffing methods, which rely on static shift schedules, fail to adapt to real-time demand, resulting in either understaffing during peak hours or inefficient use of resources during off-peak periods.

The goal of our solution was to build a forecasting-based tool that would allow for proactive, data-driven planning of staff shifts, equipment allocation, and operational prioritization.

3.2 The Forecasting Modeles

Model 1: Forecasting Daily Request Volume by Department and Task Type

Objective:
The goal of this model is to forecast the daily number of service requests, segmented by department and request type. It is designed to support general daily load planning at the departmental level and across task categories.

Model Type:
We used the XGBoost Regressor, a tree-based regression algorithm known for its high performance in modeling non-linear relationships and handling categorical variables effectively.

Features Used:

- dept_encoded: Encoded department
- req_encoded: Encoded request type
- day_encoded: Encoded day of the week (Sunday to Saturday)
- is_weekend: Binary indicator for weekend
- lag_1: Number of requests on the previous day (a time-based lag feature)

At this stage, the model was developed to provide daily forecasts of expected request volumes, per department and task type, based on historical behavioral patterns.
The model aims to serve as a tool for workload management, resource planning, and service readiness across departments in the Emergency Department.

Problem Addressed:

- Forecasting daily request volumes enables advanced planning of workforce, equipment, and task prioritization.
- For example, if a spike in medical consultation requests is predicted in Department A on Monday, staffing and preparation can be adjusted accordingly.

Algorithm Selection –XGBoost

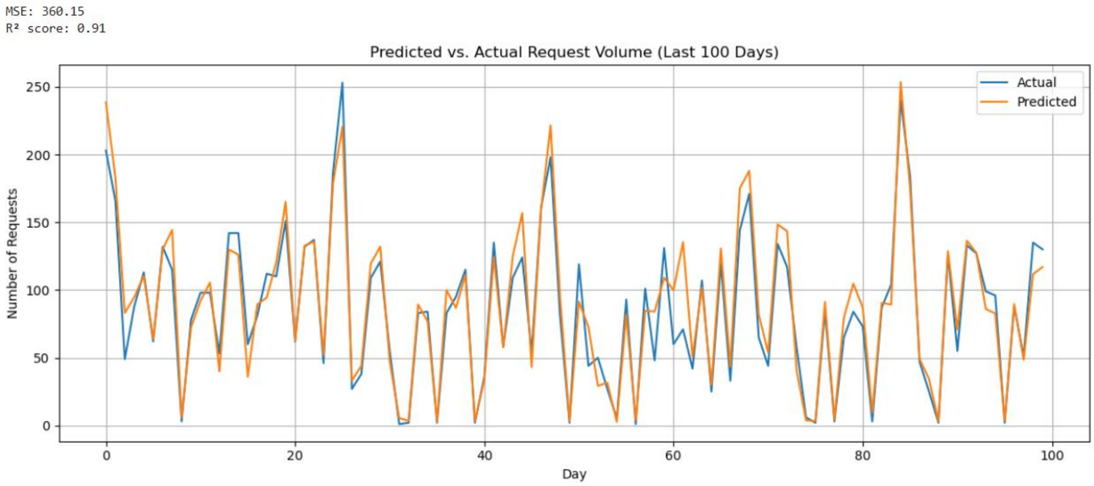
To solve the forecasting problem, we selected XGBoost (Extreme Gradient Boosting) – a powerful and widely used machine learning algorithm for supervised regression tasks.

- XGBoost is based on ensembles of decision trees that are trained sequentially, where each tree learns to correct the errors of its predecessor (a process known as boosting).
- The algorithm incorporates regularization techniques to prevent overfitting, making it well-suited for complex datasets with many interacting variables.
- It performs exceptionally well on structured data, especially when both categorical and numerical features are present – as is the case in our dataset.

Why XGBoost is a good fit for this use case:

- The data includes categorical variables (e.g., day of week, department, request type) and numerical inputs (e.g., request counts, weekend flag).
- There are non-linear interactions between features – such as the relationship between request type and weekend load.
- XGBoost effectively identifies such non-linear patterns and captures their impact on request volume.

Model Performance and Results



- R^2 Score: 0.91, indicating a very strong correlation between predicted and actual request volumes.
- Mean Squared Error (MSE): 360.15, reflecting relatively low average error given the scale of the data.

The model demonstrated excellent fit between the forecasted values and actual trends, including the ability to capture load peaks and temporal patterns.
Although some pointwise deviations were observed, the model successfully captured the overall behavioral structure of daily request volumes.

Model 2: Daily Request Load Forecast Based on Day of Week

Objective:
This model was designed to predict the total number of daily service requests in the Emergency Department, based solely on temporal features such as the day of the week, weekend indicator, and previous day's load.
The goal was to identify broad temporal patterns—independent of department or task type—and to provide a baseline forecast that relies only on time-based features.

This is a relatively simple model whose purpose is to test the feasibility of making accurate predictions without using categorical identifiers (e.g., department names or request types), relying purely on time-related variables.

Algorithm:
We used the XGBoost Regressor, a boosted decision tree algorithm highly effective for regression problems involving complex, non-linear patterns.
As in the first model, XGBoost excels in identifying interactions between features such as day of the week and request volume, even when no categorical task information is provided.

Features Used:

- day_encoded: Numerical encoding of the weekday (Sunday–Saturday)
- is_weekend: Boolean flag indicating whether the day is Friday or Saturday
- lag_1: Number of requests recorded on the previous day

Workflow:

1. Data was aggregated daily by date, combining all departments and request types.
2. All categorical variables were removed; only time-based features were retained.
3. A lag feature (lag_1) was added to reflect possible continuation or accumulation of demand.
4. Data was split chronologically: 80% of the earliest days for training, 20% for testing.
5. The model was trained and evaluated on both sets.
6. Forecasts were visualized against actual values for a 100-day horizon.

Model Performance:

Training Set:

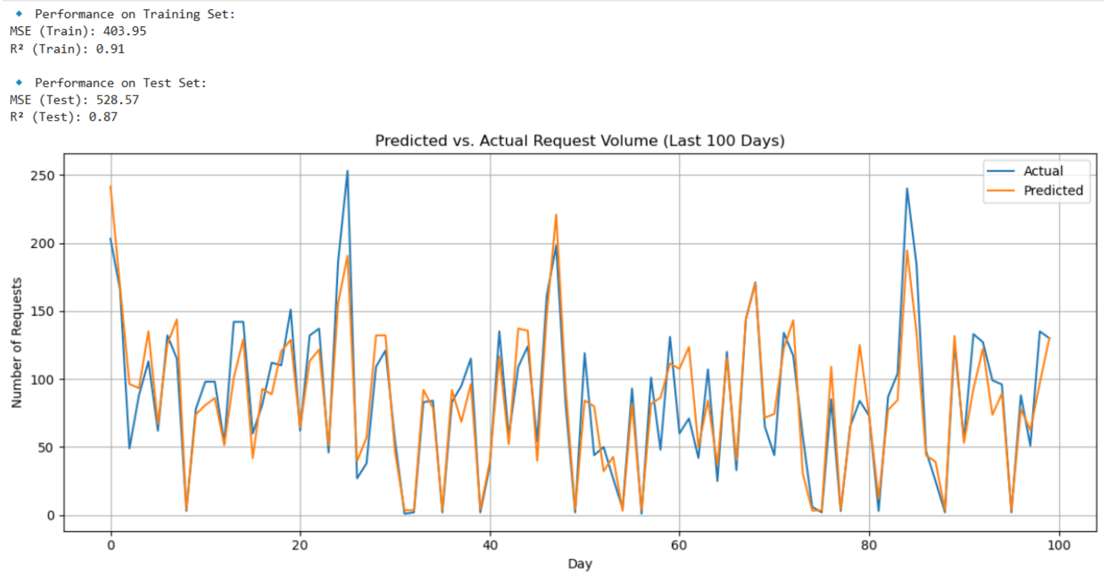
- $R^2 = 0.91$ (explaining 91% of variance)
- MSE = 403.95

Test Set (Unseen Data):

- $R^2 = 0.87$ (retaining strong accuracy)

- MSE = 528.57

The forecast line (orange) closely follows the actual data (blue), capturing **peaks and troughs** effectively.
While test error is slightly higher than on the training set—as expected—the model shows robust generalization and strong predictive capability.



Model 7-Day Daily Forecast for Each Department–Request Combination

Approach:

- A separate XGBoost model is trained for each unique (department, request type) pair.
- Each model learns the relationship between the day of the week, weekend indicator, and the previous day's request volume.
- A 7-day forecast is generated in a loop, using the output from each day as the input (lag) for the next.

Objective:
This model aims to provide a daily forecast for the upcoming 7 days for every department–request pair. It serves as an operational planning tool that enables advance preparation for expected workloads, tailored to each department and task type, based on historical patterns, weekly cycles, and weekend effects.

Technology and Algorithm:
The model uses the XGBoost Regressor, a state-of-the-art regression algorithm that leverages gradient-boosted decision trees.
It is especially well-suited for time-dependent data with recurring behavioral patterns, offering high accuracy in numeric value prediction.

Model Workflow:

1. Data Preparation
   - The raw dataset is structured by date, day of week, and a weekend flag.
   - Data is grouped by day, for each unique department–request combination.
2. Temporal Feature Encoding

- The day of the week is encoded numerically using label encoding, to allow the model to interpret cyclical time features.

3. Lag Feature Engineering
    - Each row includes lag_1, the number of requests from the previous day – a critical input for short-term prediction.
4. Model Training per Combination
    - A dedicated model is trained for each (department, request) pair with at least 10 historical observations.
    - Input features:
        - day_encoded: Numerical day of the week
        - is_weekend: Weekend indicator
        - lag_1: Previous day's request volume
5. 7-Day Forecasting
    - For each pair, the model generates a 7-day forecast using an auto-regressive method: the predicted value from one day is used as the lag for the next day.
    - The final output is stored in a structured table containing:
        - Forecast date
        - Department
        - Request type
        - Predicted request volume

Model Strengths:

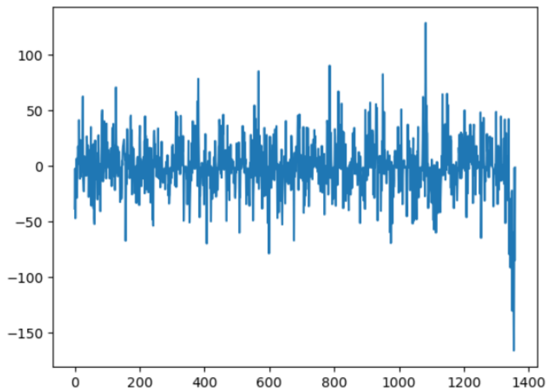
- Modular: A separate model is trained for each pair, preventing cross-contamination of behavioral patterns.
- Localized: Well-suited for capturing unique trends in specific department–task combinations.
- Realistic forecasting: Gradual predictions based on lag features ensure consistency with historical trends.

	date	department	requirement	predicted_requests
343	2024-08-21	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	2
344	2024-08-22	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	1
345	2024-08-23	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	3
346	2024-08-24	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	5
347	2024-08-25	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	2
348	2024-08-26	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	5
349	2024-08-27	מלר"ד אגף מהלכים	כח עזר - סיוע למטופל	2
280	2024-09-23	טריאז' מיון חדש	כח עזר - שונות	2
281	2024-09-24	טריאז' מיון חדש	כח עזר - שונות	3
282	2024-09-25	טריאז' מיון חדש	כח עזר - שונות	3
336	2024-09-25	מלר"ד אגף מהלכים	כח עזר - אוכל	3
283	2024-09-26	טריאז' מיון חדש	כח עזר - שונות	2
337	2024-09-26	מלר"ד אגף מהלכים	כח עזר - אוכל	2
284	2024-09-27	טריאז' מיון חדש	כח עזר - שונות	2
338	2024-09-27	מלר"ד אגף מהלכים	כח עזר - אוכל	2
285	2024-09-28	טריאז' מיון חדש	כח עזר - שונות	2

Residuals Analysis

- The analysis of residuals (i.e., the difference between predicted and actual values) shows uniform noise with no apparent structure or trend.
- There is no accumulation of errors or significant deviations, which indicates a stable and reliable model.

Residuals Plot:
The scatter plot of residuals displays a random distribution centered around zero, suggesting that there is no systematic bias in the model's predictions.

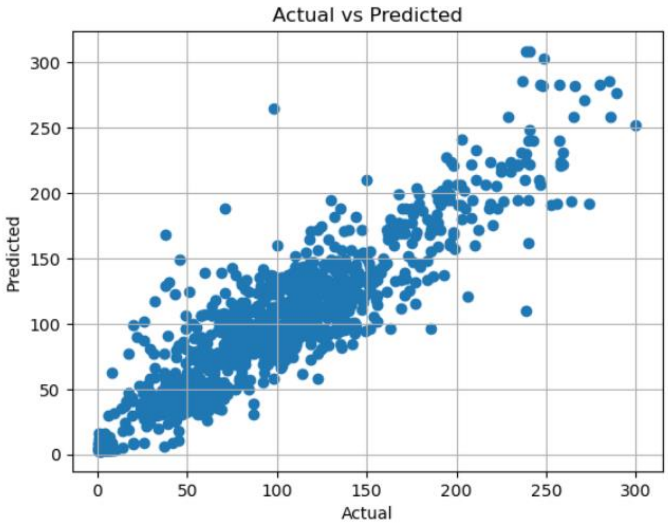


Actual vs. Predicted Comparison

- Most data points lie close to the ideal line y = x, indicating that the predicted values closely match the actual values.
- There is no evident trend of systematic under- or over-prediction.

- The model successfully captures even higher-than-usual request volumes, demonstrating robustness in edge cases.

Actual vs. Predicted Plot:
The plot exhibits strong alignment between predicted and actual values, including at the extremes, further validating the model's accuracy.



Model 3: Staffing Recommendation Model Based on Hourly Load and Task Type

Objective:
This model was developed to recommend the number of staff members needed per hour throughout the day, based on real-time request volumes.
It is designed to support shift planning and workforce scheduling in alignment with operational demand.

Approach:

- The model is based on the analysis of hourly load patterns for the six most frequent request types (referred to as "Top Tasks"), across all days of the week.
- It calculates the total request load per hour and translates it into staffing recommendations using a simple operational rule:

One staff member per 20 requests per hour (with a minimum of one staff member assigned regardless of demand level).

Goal:
The goal is to provide an initial recommendation for smart workforce allocation based on actual hourly activity patterns derived from historical data.
This allows for more efficient resource use and improved responsiveness in departments with varying load levels.

Workflow:

1. Data Preparation
    - The model reads the request dataset and extracts the hour and day of the week from each timestamp.
    - It filters the data to include only the six most frequent request types (top_tasks), in order to focus on the key contributors to overall workload.
2. Aggregation by Day and Hour
    - Data is grouped by requirement, weekday, and hour to calculate the number of requests per task per hour.
3. Load Table Construction
    - A pivot table is created, where each row represents a specific hour on a given weekday, and columns show the number of requests for each top task.
    - A total_load column is computed to represent the sum of all request types for each hour.
4. Staffing Recommendation Calculation
    - Based on the total_load, the model applies the default rule:

1 staff member per 20 requests

- The recommendation is rounded up, and a minimum of one staff member is always assigned to ensure coverage during low-demand periods.
- The resulting recommended_staff value can be used to guide shift scheduling and task distribution.

requirement	weekday	hour	בדיקת אק"ג	בדיקת דם	בדיקת רופא	ניקיון تنظيف	רמת דחיפות	תפוס	total_load	recommended_staff
0	Sunday	0.0	268.0	628.0	423.0	465.0	0.0	411.0	2195.0	110
1	Sunday	1.0	150.0	487.0	301.0	315.0	0.0	300.0	1553.0	78
2	Sunday	2.0	166.0	463.0	301.0	292.0	0.0	285.0	1507.0	75
3	Sunday	3.0	158.0	409.0	262.0	239.0	0.0	255.0	1323.0	66
4	Sunday	4.0	155.0	417.0	259.0	296.0	0.0	249.0	1376.0	69
5	Sunday	5.0	151.0	398.0	275.0	288.0	0.0	245.0	1357.0	68
6	Sunday	6.0	172.0	500.0	327.0	315.0	0.0	293.0	1607.0	80
7	Sunday	7.0	248.0	754.0	433.0	314.0	0.0	376.0	2125.0	106
8	Sunday	8.0	301.0	919.0	498.0	278.0	0.0	431.0	2427.0	121
9	Sunday	9.0	323.0	1066.0	633.0	358.0	0.0	444.0	2824.0	141
10	Sunday	10.0	319.0	1110.0	551.0	500.0	0.0	475.0	2955.0	148
11	Sunday	11.0	380.0	1122.0	570.0	455.0	0.0	490.0	3017.0	151
12	Sunday	12.0	369.0	1031.0	495.0	641.0	0.0	427.0	2963.0	148
13	Sunday	13.0	390.0	977.0	537.0	696.0	0.0	475.0	3075.0	154
14	Sunday	14.0	383.0	1011.0	604.0	678.0	0.0	509.0	3185.0	159
15	Sunday	15.0	352.0	927.0	505.0	573.0	0.0	453.0	2810.0	140
16	Sunday	16.0	320.0	931.0	539.0	496.0	0.0	442.0	2728.0	136

Chart 1: Heatmap – Request Volume by Hour and Task Type (Sunday)

What the Chart Shows:

- Each row represents a popular task type.
- Each column corresponds to an hour of the day (0–23).
- Colors and numbers indicate the number of requests submitted during that hour.

Key Insights:

1. Peak Activity Hours:
   - The task type "Sod Tikdev" (likely a cleaning or operational task) stands out with over 1,000 requests during the hours of 10:00–13:00.
   - Other tasks, such as "Apoll Tikdev", also show significant increases in the same time window.
2. Low Activity Hours:
   - Between midnight and 6:00 AM, the number of requests remains low across all task types.
3. Recurring Daily Pattern:
   - The chart illustrates a clear and consistent trend:
     - Morning = High workload
     - Nighttime = Low activity
   - This pattern appears to repeat reliably and may inform shift scheduling.

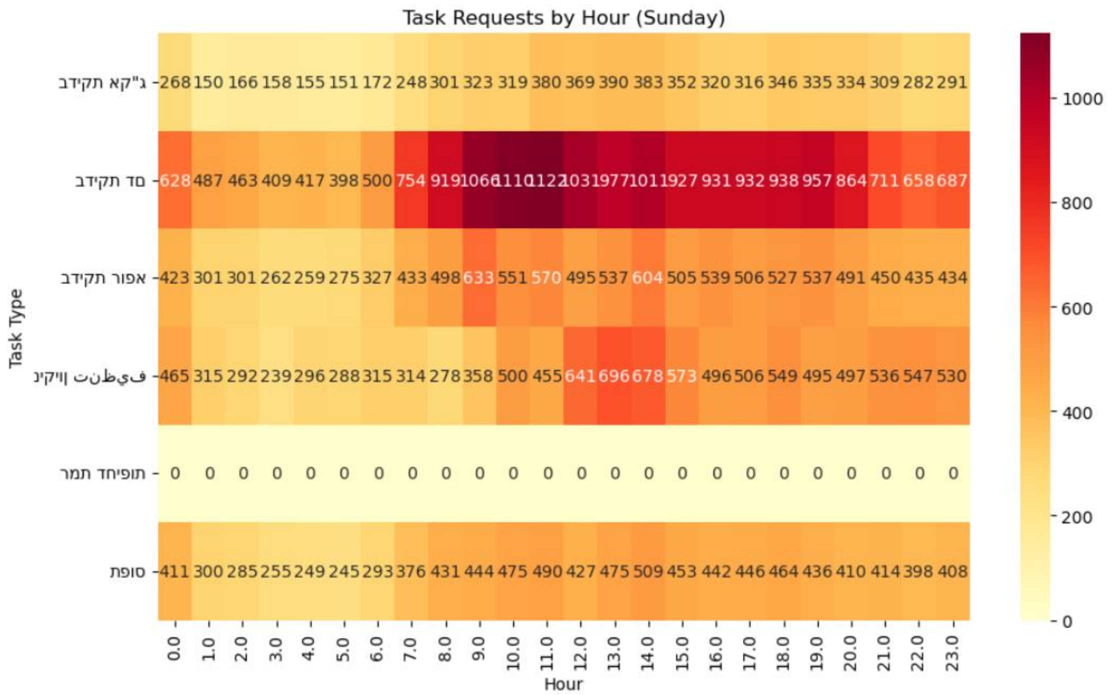


Chart 2: Total Request Volume by Hour (Sunday)

What the Chart Shows:

- A single orange line representing the total number of requests (all task types combined) per hour throughout Sunday.

Key Insights:

1. Sharp Rise in the Morning:
    - A significant increase in activity occurs between 6:00 AM and 10:00 AM, likely reflecting the start of the hospital's operational day.
2. Peak Load Hours:
    - Between 10:00 AM and 2:00 PM, request volume reaches its maximum, exceeding 3,000 requests per hour.
    - This period marks the core operational workload.
3. Gradual Decline:
    - After 3:00 PM, the request volume gradually decreases, eventually stabilizing at a lower but steady level during evening hours.

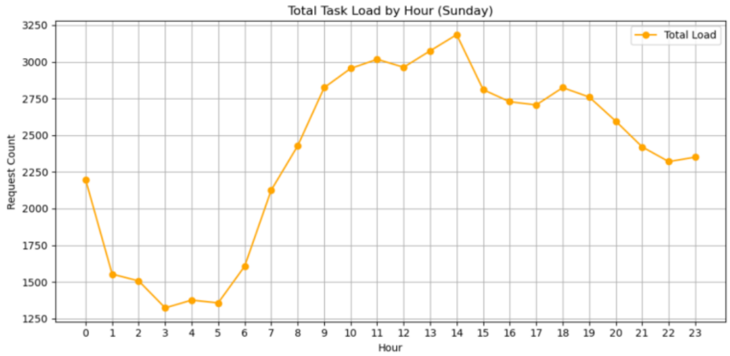


Chart 3: Recommended Number of Staff per Hour (Sunday)

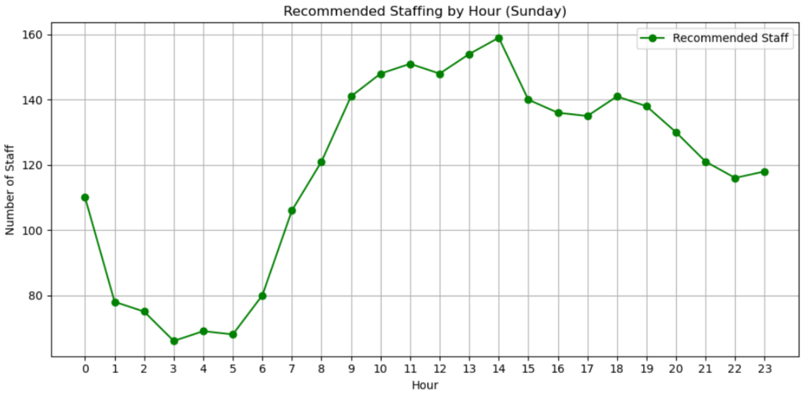
What the Chart Shows:

- A green line graph displaying the recommended number of staff members per hour, based on the total number of requests.
- The staffing is calculated using a rule of one staff member per 20 requests, with a minimum of one staff member per hour.

Key Insights:

1. Peak Staffing Hours:
    - Between 10:00 AM and 2:00 PM, the model recommends assigning 145 to 155 staff members.
    - This corresponds directly with the high task volume observed during the same period.

2. Minimum Staffing at Night:
   - During early morning hours (2:00–6:00 AM), the recommendation drops to 65–75 staff members, reflecting reduced demand.
3. Responsive to Demand Patterns:
   - The staffing recommendation curve closely follows the request load trends seen in Chart 2, indicating that the model appropriately adjusts to workload fluctuations throughout the day.



3.3 Summary of Key Insights from the Results

Practical Recommendations

1. Design shift schedules based on actual load patterns, rather than fixed time blocks.
2. Assign staff according to task types – for example, designate focused teams for high-demand tasks during peak hours.
3. Replicate this analysis for other weekdays to determine whether Sunday’s pattern is consistent across the week.
4. Refine the “20 requests per staff member” rule by incorporating real-time task duration data in future versions of the model.