**IMAT-5264**

**REPORT ON DATA ANALYSIS ON NHS USING EXCEL**



*Done by*

Aathira Puthenpurayil, P2736418

Jackson Kalathil George, P2725132

Juny Maria John, P2743266

Navya Anjinappa, P2729195

Chippy Sundaran Pillai, P2758512

*Under the guidance of*

**Mr Geoff Payne**

*Professor at the Department of Computing, Engineering & Media*

**Master of Science**

**In**

**Data Analytics**

**Faculty of Computing, Engineering & Media (CEM)**

**De Montfort University**

**Leicester – LE1 9BH**

Table of Contents

[**Introduction** 2](#_Toc135255887)

[**Forecasting Evaluation** 2](#_Toc135255888)

[**Data Access, Cleaning and Preparation** 3](#_Toc135255889)

[**Cancer Waiting Times with a 62-day waiting standard** 3](#_Toc135255890)

[**Beds by the Board of Treatment and Speciality** 5](#_Toc135255891)

[**Time Series Forecasting** 6](#_Toc135255892)

[**Linear Optimisation** 7](#_Toc135255893)

[**Ethical Issues in Business Intelligence** 8](#_Toc135255894)

[**References** 9](#_Toc135255895)

[**Appendices** 10](#_Toc135255896)

# **Introduction**

The National Health Service's (NHS) effectiveness can be evaluated by taking into account the length of the cancer treatment waitlist and the occupancy rate of NHS hospitals. These factors offer helpful perceptions and insights into the effectiveness of the healthcare system, allowing medical professionals to identify areas that need to improve patient care quality.

To ensure that patients receive quick treatment, which is crucial for their survival and recovery, patients with suspected cancer have a maximum 62-day wait period before receiving their first cancer treatment following an urgent referral from their GP. The NHS bed occupancy rate refers to the percentage of hospital beds that are currently occupied by patients to assess the NHS's capacity to handle the demand for medical care. Increased bed occupancy can lead to overcrowding, protracted waiting times, and a higher risk of infection, all of which can be detrimental to patients’ well-being.

The information about cancer waiting times and bed occupancy forecasting using linear trend, the linear trend with seasonality, and exponential smoothing is examined in this report to help medical professionals streamline resource distribution, control patient traffic, increase patient satisfaction, and reduce operational costs while maintaining or raising the standard of medical care.

# **Forecasting Evaluation**

1. Forecasting techniques play a critical role in ensuring that NHS Scotland allocates resources effectively and meets the healthcare needs of the population. To measure the success of their forecasting techniques, management at NHS Scotland should consider four key questions.

Firstly, they should assess whether they have built an unbiased momentum case that incorporates internal and external data, and end-market trends, and allows for new information to be layered on top.

Secondly, they should determine whether they are using a variety of operational indicators and external inputs, such as hospital admissions and waiting times, to better anticipate changes in demand for healthcare services.

Thirdly, companies should assess if they are utilising technology, such as sophisticated analytics tools and artificial intelligence, to provide more precise projections and make better resource allocation decisions.

To guarantee that everyone is working towards the same goals and has access to the most recent data and insights, they should evaluate how well they are collaborating across departments.

1. However, there are several challenges and implications that NHS Scotland may face when applying forecasting approaches and processes.
2. Data quality is crucial for accurate forecasts, and incomplete or inaccurate data can lead to inaccurate forecasts.
3. Data privacy and security concerns may arise when sharing patient data across departments or with external partners.
4. Resistance to change and reluctance to adopt new forecasting approaches could lead to continued use of outdated methods.
5. Lack of expertise in the healthcare industry and advanced analytics could affect the accuracy of forecasts.
6. Funding constraints could limit investment in technology and resources required for accurate forecasting.

# **Data Access, Cleaning and Preparation**

# **Cancer Waiting Times with a 62-day waiting standard**

The plot represents the percentage of referrals for cancer that have received treatment in Scotland NHS from 2012 to 2021.

Over time, more recommendations have been accepted for treatment on average. However, the data indicate that there are significant differences in the treatment rates across the NHS, with some achieving rates close to 100% and others falling below 80%. The percentages of referrals treated in 2014 by NHS Western Isles (66.66%) and The Golden Jubilee National Hospital (55.55%) both experienced notable drops.

**Chart, line chart

Description automatically generated**

The covid pandemic struck a portion of the NHS in the year 2020 (period 33–36), which results in a decline in the proportion of referrals treated. This can be mainly due to the high demand for intensive care beds during that period [2]. The dips in the plot show the challenges in cancer treatment which have been a persistent issue in Scotland. For example, a 2021 article in BBC News reported that a cancer patient had to cancel her surgery three times in a month due to a lack of intensive care beds at Ninewells Hospital in Dundee. The article also shows that 407 planned operations were cancelled the day before or on the day due to capacity or non-clinical reasons [1]. To further investigate the percentage of referrals that received treatment, it is essential to consider other factors that may influence treatment rates, hospital size, location, resources etc. It is also crucial for hospitals to adopt efficient & effective referral and treatment protocols to reduce waiting time and improve treatment rates [3].

# **Beds by the Board of Treatment and Speciality**

A graph of a bed occupancy

Description automatically generated with low confidence

The plot shows there are significant bed shortages in the Medical Oncology department in Scotland. But NHS Tayside got a 100% of bed occupancy rate throughout the given period. From the graph, it is also noticed that all Health boards except the Tayside have a huge in-bed occupancy rate during the pandemic period. But the health board including NHS Lothian, Scotland, NHS Grampian, and NHS Greater Glasgow and Clyde bounces back to their initial occupancy percentages.

The article ‘Bed blocking is crippling Scotland’s NHS’ [5] referred to some of the challenges of bed shortages in the Medical Oncology department in Scotland due to various factors such as the continuous increase in the number of cancer patients which made the hospitals allocate all of them. Another reason found is due to the highly old age population as they need quick treatment with proper care. In addition to it, the article mentions the shortage of staff and strikes going on in many NHS which is a great challenge for patients’ treatment [1]. Furthermore, COVID Pandemic [4] has also been a major contributor as the demand for beds was very high and the health boards in Scotland don’t have the required number of beds for the treatment.

# **Time Series Forecasting**

A picture containing text, screenshot, font, number

Description automatically generated

A picture containing text, screenshot, font, number

Description automatically generated

As the data has a constant mean and the statistical properties are independent of time, the percentage occupancy of hospital beds for the Tayside health board from 2018 to 2021 follows a stationary time series. This also applies to Tayside data without 2020. The horizontal pattern and random variations that the time series plot displays are proof that it is stationary. Due to its consistent data and lower Mean Squared Error than other approaches, Tayside's cancer referral data uses exponential smoothing with smoothing constant α = 0.7 when including and omitting 2020.

Because of its general stability, lack of trend or seasonality, and lower Mean Squared Error under exponential smoothing with a smoothing constant of = 0.9 than under linear trend and linear trend with seasonality, the Lothian bed occupancy data is thought to follow a horizontal pattern. However, when we excluded 2020 from the analysis, we could find a linear trend with seasonality (Mean Squared Error is smaller than other approaches). Lothian's cancer referral data use exponential smoothing with a smoothing constant of 0.8 when including and omitting the year 2020 due to its consistent data and lower Mean Squared Error than other techniques.

Since the time series exhibit both trend and seasonality, we could see that the hospital beds data in Grampian follows a linear trend with seasonality while including and excluding the year 2020. Because it has less Mean Squared Error than other methods, the cancer referral data in Grampian follows a linear trend with seasonality while including and excluding the year 2020.

# **Linear Optimisation**

**Decision Variables**

* X1 = Number of staff for 09.00-13.00
* X2 = Number of staff for 13.00-17.00
* X3 = Number of staff for 17.00-21.00
* X4 = Number of staff for 21.00-01.00
* X5 = Number of staff for 01.00-05.00
* X6 = Number of staff for 05.00-09.00

**Result Variable and Objective**

The objective is to minimize the total number of staff required for 24 hours.

* Z = X1 + X2 + X3 + X4 + X5 + X6

**Uncontrollable variables (Constraints)**

* X1 + X2 + X3 >= 36 (for shift 1)
* X2 + X3 + X4 >= 26 (for shift 2)
* X3 + X4 + X5 >= 22 (for shift 3)
* X4 + X5 + X6 >= 16 (for shift 4)
* X5 + X6 + X1 >= 10 (for shift 5)
* X6 + X1 + X2 >= 6 (for shift 6)

**Justification**

The constraint "X1 + X2 + X3 >= 36" means that the total number of staff required for shifts 1, 2, and 3 combined should be equal to or greater than 36. This ensures that there are enough staff available during the morning shift (09:00-13:00) to meet the demand of 36 staff members.

Similarly, the other constraints ensure that the demand for staff is met every 4 hours of the day.

# **Ethical Issues in Business Intelligence**

The use of business intelligence (BI) technologies in healthcare, particularly in NHS Scotland, raises significant ethical, legal, and privacy problems that could jeopardise patient information availability, confidentiality, and integrity. They must address critical problems such as patient privacy, informed consent, bias and discrimination, openness and accountability, and data security when using BI systems.

**Patient privacy** [6]is a critical concern and must ensure that data is encrypted, access is restricted to approved workers only, and data breaches are reported to the Information Commissioner's Office (ICO) if the BI system collects information that can be used to identify specific patients. Informed consent is also essential, and NHS Scotland must tell patients about data collection and planned uses, as patients have the right to know how their data is being used and to opt-out.

Business intelligence systems can be **biased and discriminatory** [6], thus NHS Scotland must verify that the algorithms used to train the system are fair and that no patients are singled out for prejudice. **Transparency and accountability** [6] are also critical and must be upfront and honest about how data is used, who has access to it, and if it can be confirmed.

Finally, they must **protect data** [6]from unlawful access, alteration, or destruction. If the BI system is used to identify patients at high risk of getting a specific disease, the data must be securely maintained, and only approved workers should be provided access.

To summarise, implementing BI tools in NHS Scotland requires careful consideration of ethical, legal, and privacy concerns to guarantee that patient data is properly handled and used. Following these guidelines will increase trust and confidence in the usage of BI technologies in healthcare.

# **References**

1. Cancer patient’s surgery was cancelled three times due to ICU beds shortage. (2021). *BBC News*. [online] 7 Dec. Available at: <https://www.bbc.co.uk/news/uk-scotland-tayside-central-59561747>.
2. Scottish cancer waiting times performance at record low. (2022). *BBC News*. [online] 28 Jun. Available at: <https://www.bbc.co.uk/news/uk-scotland-61968630>.
3. Covid in Scotland: Thousands fewer had cancer treatment. (2021). *BBC News*. [online] 29 Jun. Available at: <https://www.bbc.co.uk/news/uk-scotland-57650825>.
4. Scottish hospitals are almost full, says Nicola Sturgeon. (2023). *BBC News*. [online] 9 Jan. Available at: <https://www.bbc.co.uk/news/uk-scotland-scotland-politics-64211393>.
5. bed-blocking-is-crippling-Scotland’s-nhs ,says [Lucy Dunn](https://www.spectator.co.uk/writer/lucy-dunn/).(2023).*The Spectator*. [online] 13 Jan. Available at: <https://www.spectator.co.uk/article/bed-blocking-is-crippling-scotlands-nhs/>
6. Guidance data ethics framework, GOV.UK(2020) [online] 16 Sep. Available at: <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020>

# **Appendices**

**Management Evaluation Report**

|  |  |  |  |
| --- | --- | --- | --- |
| Student ID | P2736418 | Contribution  %\* | 20% |
| Justification | I, cleaned beds by board treatment and specialty and also did time series forecasting for Lothian(S08000024) in 2020. Helped the team with Linear optimization. | | |
| Student ID | P2725132 | Contribution  %\* | 20% |
| Justification | Jackson cleaned the Cancer waiting Times with a 62-day waiting standard and did the forecasting evaluation. He has also contributed to forecasting Grampian(S08000020) with 2020 and helped the team with Linear optimization. | | |
| Student ID | P2729195 | Contribution  %\* | 20% |
| Justification | Navya has done time series forecasting for Tayside(S08000030) with 2020. She had done ethical issues in Business intelligence and helped the team with Linear Optimization. | | |
| Student ID | P2758512 | Student ID  %\* | 20% |
| Justification | Chippy contributed towards Time series forecasting for Lothian(S08000024) without 2020 and time series forecasting for Tayside(S08000030) without 2020. She has also done observations on the time series and helped the team with Linear optimization. | | |
| Student ID | P2743266 | Contribution  %\* | 20% |
| Justification | Juny cleaned beds by the board of treatment and specialty and also did time series forecasting for Grampian without 2020. She also helped the team with linear optimization. | | |

**Your student Id: P2736418**

|  |  |  |  |
| --- | --- | --- | --- |
| Student ID | P2743266 | Contribution  %\* | 20% |
| Justification | Cleaned data from beds by the board of treatment and specialty CSV file and did the forecasting. Also did time series forecasting for Grampian without 2020. Helped the team with linear optimization. | | |
| Student ID | P2725132 | Contribution  %\* | 20% |
| Justification | Jackson cleaned data from the Cancer waiting Times with a 62-day waiting standard and did the forecasting. He has also forecasted Grampian(S08000020) with 2020 and helped the team do the Linear optimization. | | |
| Student ID | P2729195 | Contribution  %\* | 20% |
| Justification | Navya has done time series forecasting for Tayside(S08000030) with 2020. She had done legal and ethical issues in Business intelligence and helped the team with Linear Optimization. | | |
| Student ID | P2758512 | Student ID  %\* | 20% |
| Justification | Chippy contributed towards Time series forecasting for Lothian(S08000024) and Tayside(S08000030) without 2020. She has also done observations on the time series and helped the team with Linear optimization. | | |
| Student ID | P2736418 | Contribution  %\* | 20% |
| Justification | Cleaned data from beds by the board of treatment and specialty CSV file and also did time series forecasting for Lothian(S08000024) in 2020. Helped the team with Linear optimization. | | |

**Your student Id: P2743266**

|  |  |  |  |
| --- | --- | --- | --- |
| Student ID | P2729195 | Contribution  %\* | 20% |
| Justification | I worked on time series forecasting for Tayside(S08000030) in 2020. Also worked on ethical issues in Business intelligence and helped the team with Linear Optimization. | | |
| Student ID | P2725132 | Contribution  %\* | 20% |
| Justification | Jackson cleaned the Cancer waiting Times with a 62-day waiting standard and did the forecasting evaluation. He has also contributed to forecasting Grampian(S08000020) with 2020 and helped the team with Linear optimization. | | |
| Student ID | P2736418 | Contribution  %\* | 20% |
| Justification | Aathira Cleaned beds by board treatment and specialty and also did time series forecasting for Lothian(S08000024) in 2020. Helped the team with Linear optimization. | | |
| Student ID | P2758512 | Student ID  %\* | 20% |
| Justification | Chippy contributed towards Time series forecasting for Lothian(S08000024) without 2020 and time series forecasting for Tayside(S08000030) without 2020.  She has also done observations on the time series and helped the team with Linear optimization. | | |
| Student ID | P2743266 | Contribution  %\* | 20% |
| Justification | Juny cleaned beds by the board of treatment and speciality and also did time series forecasting for Grampian without 2020. She also helped the team with linear optimization. | | |

**Your student Id: P2729195**

|  |  |  |  |
| --- | --- | --- | --- |
| Student ID | P2758512 | Contribution  %\* | 20% |
| Justification | Contributed towards Time series forecasting for Lothian(S08000024) and Tayside(S08000030) without 2020. Also done observations on the time series and helped the team with Linear optimization. | | |
| Student ID | P2725132 | Contribution  %\* | 20% |
| Justification | Jackson ensured that there was team coordination so that everyone could work effectively and efficiently to accomplish the goals. He cleaned data from the Cancer waiting Times with a 62-day waiting standard and did the forecasting. He has also forecasted Grampian(S08000020) with 2020 and helped the team do the Linear optimization. | | |
| Student ID | P2729195 | Contribution  %\* | 20% |
| Justification | Navya has done time series forecasting for Tayside(S08000030) with 2020. She had done legal and ethical issues in Business intelligence and helped the team with Linear Optimization. | | |
| Student ID | P2743266 | Student ID  %\* | 20% |
| Justification | Juny cleaned data from beds by the board of treatment and speciality CSV file and did the forecasting. Also did time series forecasting for Grampian without 2020. Helped the team with linear optimization. | | |
| Student ID | P2736418 | Contribution  %\* | 20% |
| Justification | Cleaned data from beds by the board of treatment and speciality CSV file and also did time series forecasting for Lothian(S08000024) in 2020. Helped the team with Linear optimization. | | |

**Your student Id: P2743266**

|  |  |  |  |
| --- | --- | --- | --- |
| Student ID | P2725132 | Contribution  %\* | 20% |
| Justification | I performed cleaning Cancer Waiting Times with a 62-day waiting for standard Excel. Also done Forecasting evaluation, forecasted Grampian (S08000020) with 2020 and linear optimization | | |
| Student ID | P2758512 | Contribution  %\* | 20% |
| Justification | Chippy performed time series forecasting for Lothian (S08000024) without 2020. Also, time series forecasting for Tayside (S08000030) without 2020, linear optimization and observations on the time series forecasting analysis and why chosen the specific forecasting approach | | |
| Student ID | P2729195 | Contribution  %\* | 20% |
| Justification | Navya performed time series forecasting for Tayside(S08000030) in 2020. Also, Ethical Issues in Business Intelligence and linear optimization. | | |
| Student ID | P2743266 | Student ID  %\* | 20% |
| Justification | Juny performed Cleaning beds by boards of treatment and speciality. Also, time series forecasting for Grampian without 2020 and linear optimization | | |
| Student ID | P2736418 | Contribution  %\* | 20% |
| Justification | Aathira performed cleaning beds by the board of treatment and speciality. Also, time series forecasting for Lothian (S08000024) with 2020 and linear optimization. | | |

**Your student Id: P2725132**