Analysis of E. coli and MSSA Infection Rates During COVID-19 Period

Executive Summary

This document provides technical documentation for an R script that analyzes the impact of COVID-19 on infection rates. The analysis focuses on E. coli and Methicillin-sensitive Staphylococcus aureus (MSSA) infections across different demographic groups.

Data Description

**Source File:** Tempotal\_Trends.ods (Sheet 5)

**Key Variables:**

- Infection rates by pathogen (E. coli and MSSA)

- Gender-specific rates per 100,000 population

- Age group stratification

- Financial year data points

Required R Packages

\* tidyverse

\* readxl

\* readODS

\* ggplot2

\* gridExtra

\* dplyr

\* grid

Implementation Guide

1. **Data Preparation**

The script begins by importing and preparing the data:

```R

Age\_Cases <- read\_ods("path/to/Tempotal\_Trends.ods",

sheet = 5,

skip = 2)

```

2. **Analysis Components**

The script performs several types of analyses:

**Time Series Analysis**

\* Trends over financial years

\* Period comparisons (Pre/During/Post-COVID)

\* Year-over-year changes

**Statistical Analysis**

\* ANOVA testing

\* Kruskal-Wallis analysis

\* Pairwise comparisons with Bonferroni correction

**Impact Assessment**

\* Rate changes during COVID-19

\* Recovery pattern analysis

\* Gender and age group comparisons

Customization Guide

Modifiable Parameters

1. **Time Periods**

```R

create\_time\_period <- function(year) {

if (year <= "2019 to 2020") return("Pre-COVID")

if (year >= "2022 to 2023") return("Post-COVID")

return("During-COVID")

}

```

2. **Age Groups**

```R

age\_order <- c("Fewer than 1", "1 to 14", "15 to 44",

"45 to 64", "65 to 74", "75 to 84",

"85 and over")

```

3. **Statistical Parameters**

Statistical tests can be modified in the `run\_statistical\_tests()` function.

Data Assumptions

- Consistent data format

- Complete time series

- Valid numerical entries

- No missing values in key fields – This dataset had no NA values but we have not taken out NA values because it is important to know which NA values are missing

**Statistical Considerations**

- Significance level: p < 0.05

- Bonferroni correction for multiple comparisons

- Non-parametric testing where appropriate