2025 Biennium

Projections: Traditional Medicaid

Legislative Fiscal Division

Julia Hamilton

Josh Poulette

Agenda

- General Approach
- Data
- Time Series Models
- SARIMA Models
- Uncertainty Looking Forward
- Appendix
  - Statistical Assumptions
  - Statistical Form
  - R Programming Language

General Approach

Data driven projections based on...

- Data from previous points in time
- Observed trends in expenditure

Our goal

- Optimize accuracy
- Minimize error

Data

- Medicaid paid claims data
  - Monthly data which is later aggregated to predict yearly claims values
  - 2004 – Present
  - Received from the agency on a monthly basis (901s)

- Projected individually by provider type as well as aggregated by division (BHDDD, SLTC, HRD)

- Some categories that we do not project are filled in with estimates from DPHHS

Data – Incomplete Years

New data is received for any particular service year up to 3 years after the end of that calendar year.

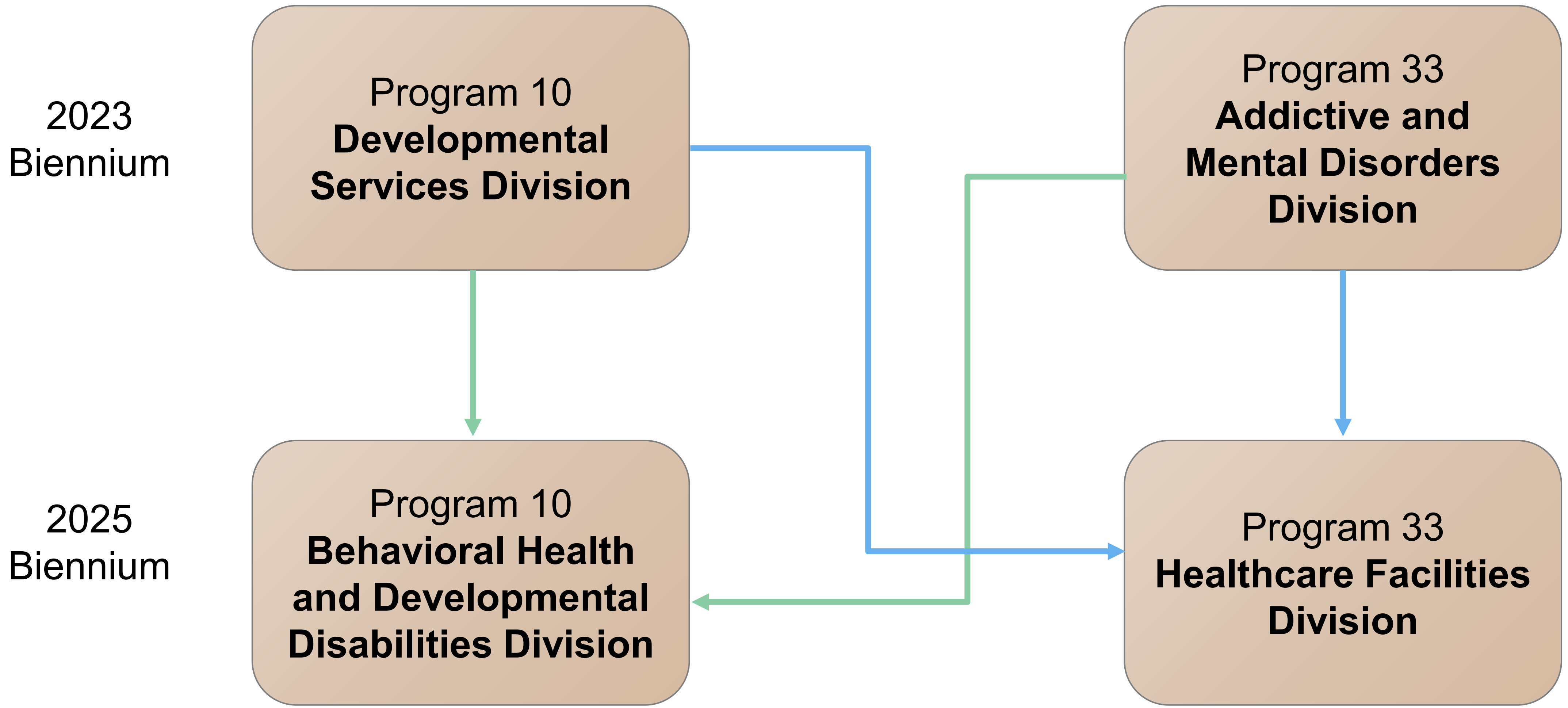
Solution: The values used for FY 2021, FY 2022, and FY 2023 are forecasted using a rate of payment completion for each service by respective month.

- These values are informed by actual payments but they are estimates and not actual values

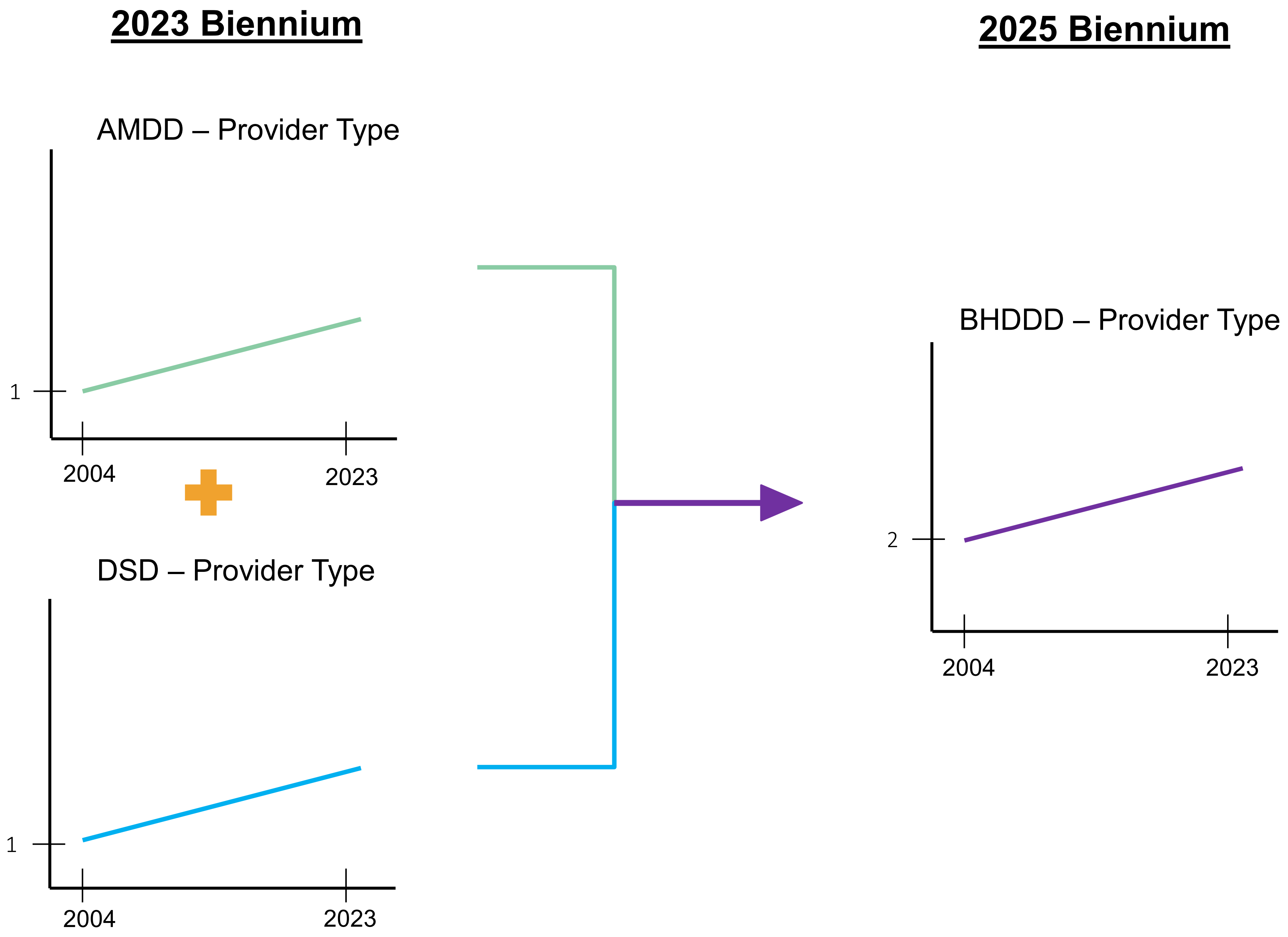
Data – The Reorganization

Non-institutional Functions

Institutional Functions



Reorganizing the Data



Time Series Data and Models

Time Series Introduction

What is a time series?

- **A set of data collected and indexed sequentially in fixed intervals**

Using time series data we can predict likely future outcomes, given knowledge of the most recent outcomes:

– More recent observations are weighted more heavily

– Implicitly accounts for inflationary and other macroeconomic factors

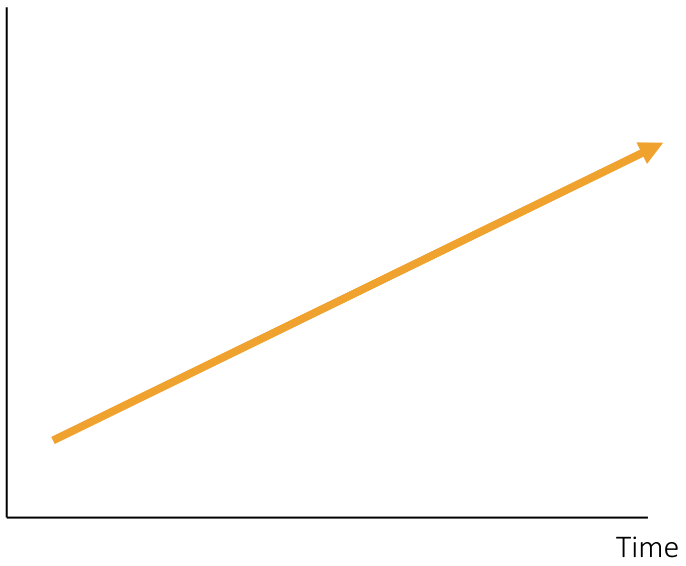
– One-time shocks (the pandemic, financial crises) affect subsequent values infinitely into the future though their explanatory power continually declines

– Cannot intrinsically adjust for policy change

Time Series Model Components

Trend

- Captures the long term
- The general tendency of data to increase or decrease over time
- General medical inflation is captured under this component



Long-term trend

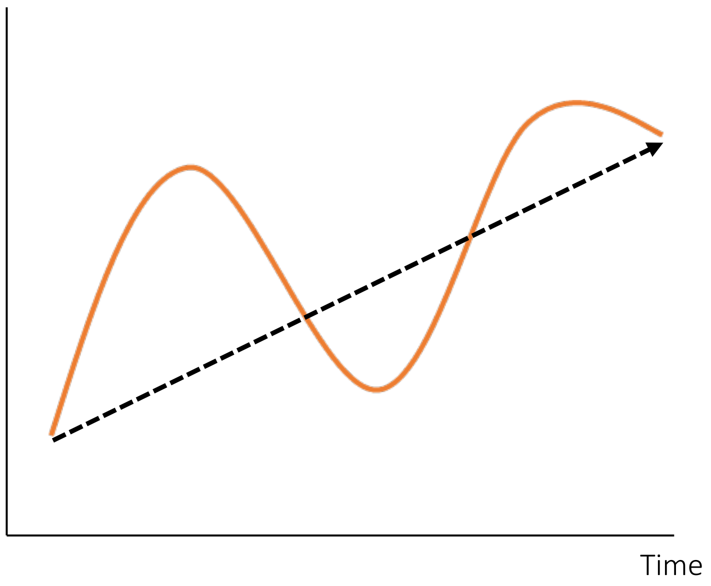
Time Series Model Components

Cyclical Factors

– Medium-term cycles, generally over 2 or more years

– Captures the parts of the data which can be explained by other cyclical movements in the economy (i.e. unemployment)

– Can occur at any time of year; unpredictable



Long-term trend with cyclical variations

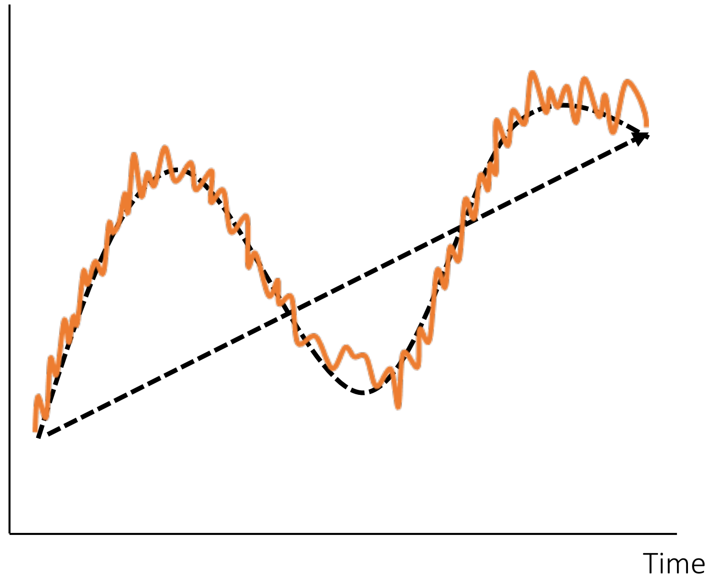
Time Series Model Components

Seasonality

– Captures short-term cycles

– The variation in a variable due to some predetermined patterns in its behavior; predictable

– Repetitive patterns that show up at certain times of the year (Ex: Increased hospital expenditures in the winter due to influenza)



Long-term trend with cyclical and seasonal variations

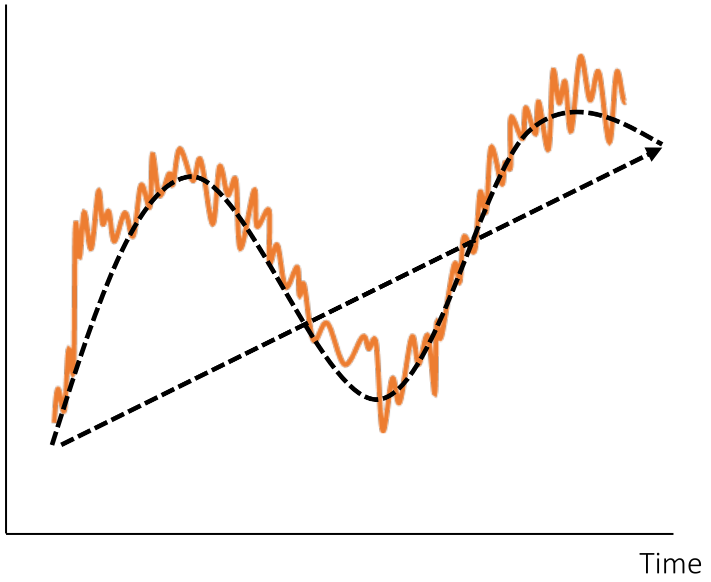
Time Series Model Components

Residual Variation

– Unpredictable random influences outside of regular patterns; it basically captures everything left over after the other three components are isolated

– Also called random noise or random variation

–Cannot be replicated by repeating an experiment again



Long-term trend with cyclical, seasonal and residual variations

Seasonal ARIMA Model

A combination of two statistical models: a generalized **auto-regressive (AR)** model **integrated (I)** with a **generalized moving average (MA)** model and **a seasonal component.**

AR(p)	Integrated (I)	MA(q)	Seasonal Component
✓ Uses previous data values to make predictions	✓ Uses differenced data to help satisfy statistical assumption of stationarity	✓ Uses the series mean and previous errors to make predictions	✓ A correction for when the data shows a regular or predictable pattern that repeats over a calendar year

For statistical form and assumptions, please visit the appendix at the end of this presentation.

Auto-Regressive

- Incorporates **cyclical** trends in the data or long-term patterns of ups and downs
- Implicitly incorporates larger macroeconomic variables

Integrated

- R software performs a mathematical transformation to satisfy the stationarity assumption, makes a prediction, and inversely transforms the result

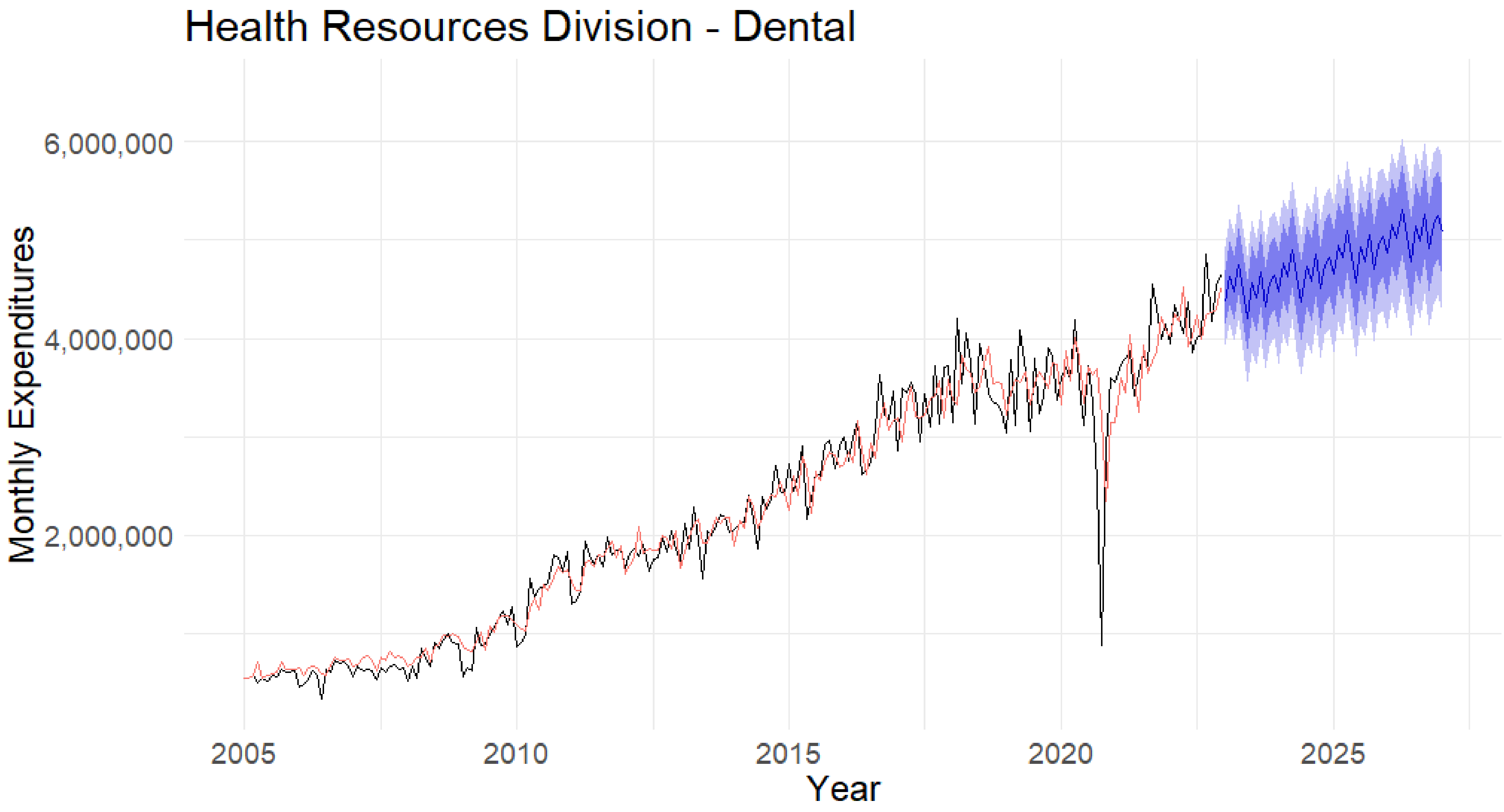
Moving Average

- Incorporates the overall **trend** and the changing of the trend over time

Seasonal Component

- Incorporates the **seasonal** or repetitive patterns that consistently show up throughout the year due to behavioral changes

Projection Example



Projection Example

Division	Provider Type	FY 2022	FY 2023	FY 2024	FY 2025
HRD	**Dental**	$51,050,982	$54,131,285	$56,046,931	$58,398,229

Uncertainty Looking Forward

Unwinding Medicaid after the Public Health Emergency (PHE)

- Disenrollment Restrictions Lifted

  Redetermination start date set to April 1, 2023

- Enhanced Federal Medical Assistance Percentage (FMAP)

SFY 2022 Rate for…	FMAP	PHE FMAP
Traditional Medicaid	64.32	70.52
Medicaid Expansion	90.00	90.00
Children’s Health Insurance Plan (CHIP)	75.02	81.22

Federal phasing out of enhanced rates

Q1 2023: 6.2 percentage point enhancement

Q2 2023: 5.0 percentage point enhancement

Q3 2023: 2.5 percentage point enhancement

Q4 2023: 1.5 percentage point enhancement

Source: https://medicaiddirectors.org/

The Big Picture

Data driven approach reliant on 19 years of monthly data that is then presented in the aggregate as an annual estimate.

Major uncertainty following the end of the PHE, which includes:

- *The reinstatement of normal disenrollment protocol with the end of the PHE*
- *Potential economic contraction, which could put upward pressure on enrollment*

Questions?

Appendix

Appendix – Statistical Assumptions

Stationarity must be satisfied.

What does this mean?

- A series whose properties do not depend on the time at which the series is observed.

Or

- The time series must have no predictable patterns in the long term

Mathematically, the mean, variance and autocovariance must remain constant over time

1. $E[y_t] = E[y_{t+k}]$
2. $var(y_t) = var(y_{t+k})$
3. $cov(y_t, y_{t+k}) = cov(y_{t+i}, y_{t+k+i})$

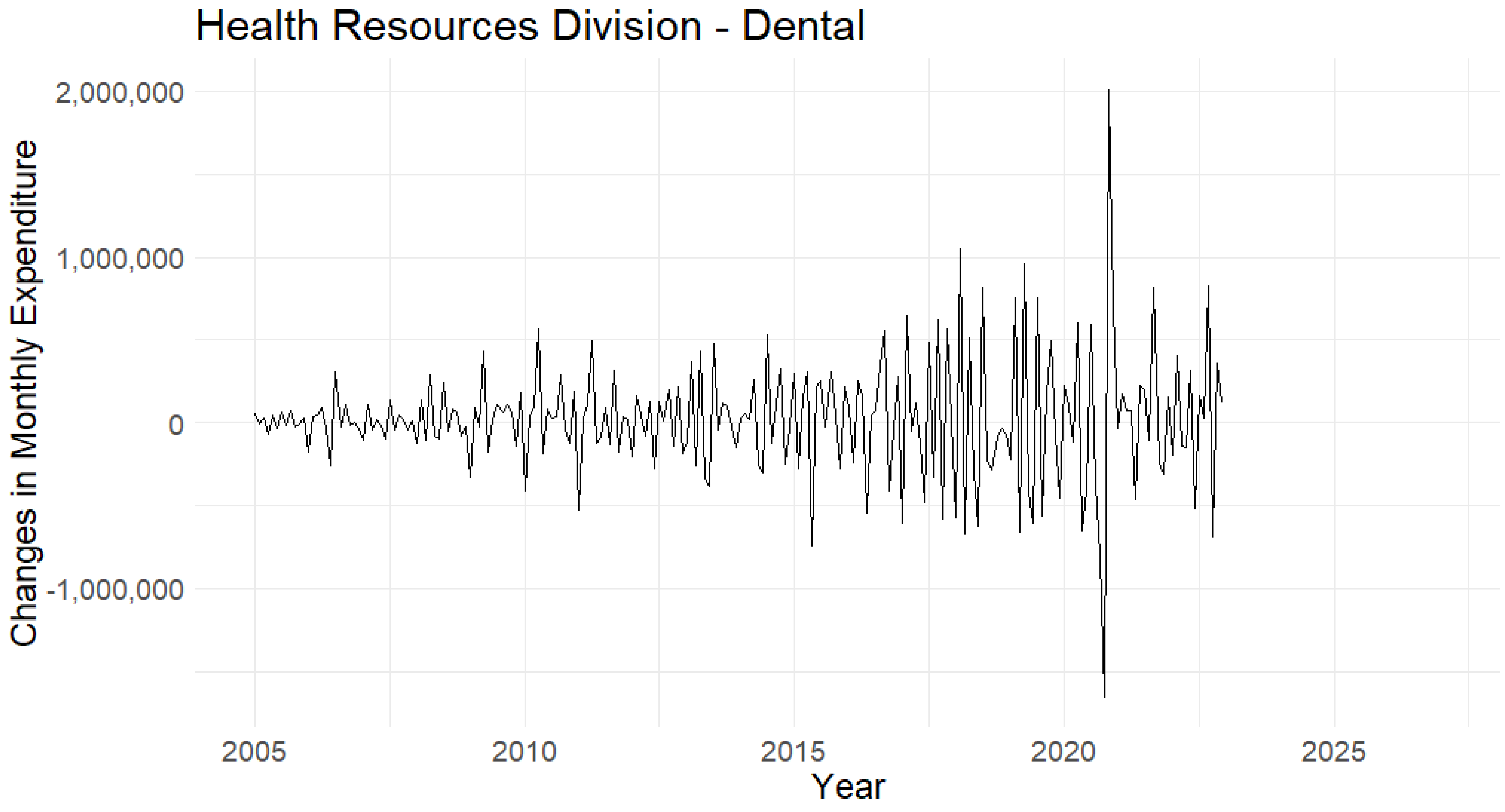
Appendix – Satisfying Stationarity

Differencing.

The most common method to satisfy the stationarity of a time series is to take the difference between one observation and the previous observation.

Mathematically…

$$y'_t = y_t - y_{t-1}$$


Health Resources Division - Dental
Changes in Monthly Expenditure
2,000,000
1,000,000
0
-1,000,000
2005
2010
2015
2020
2025
Year

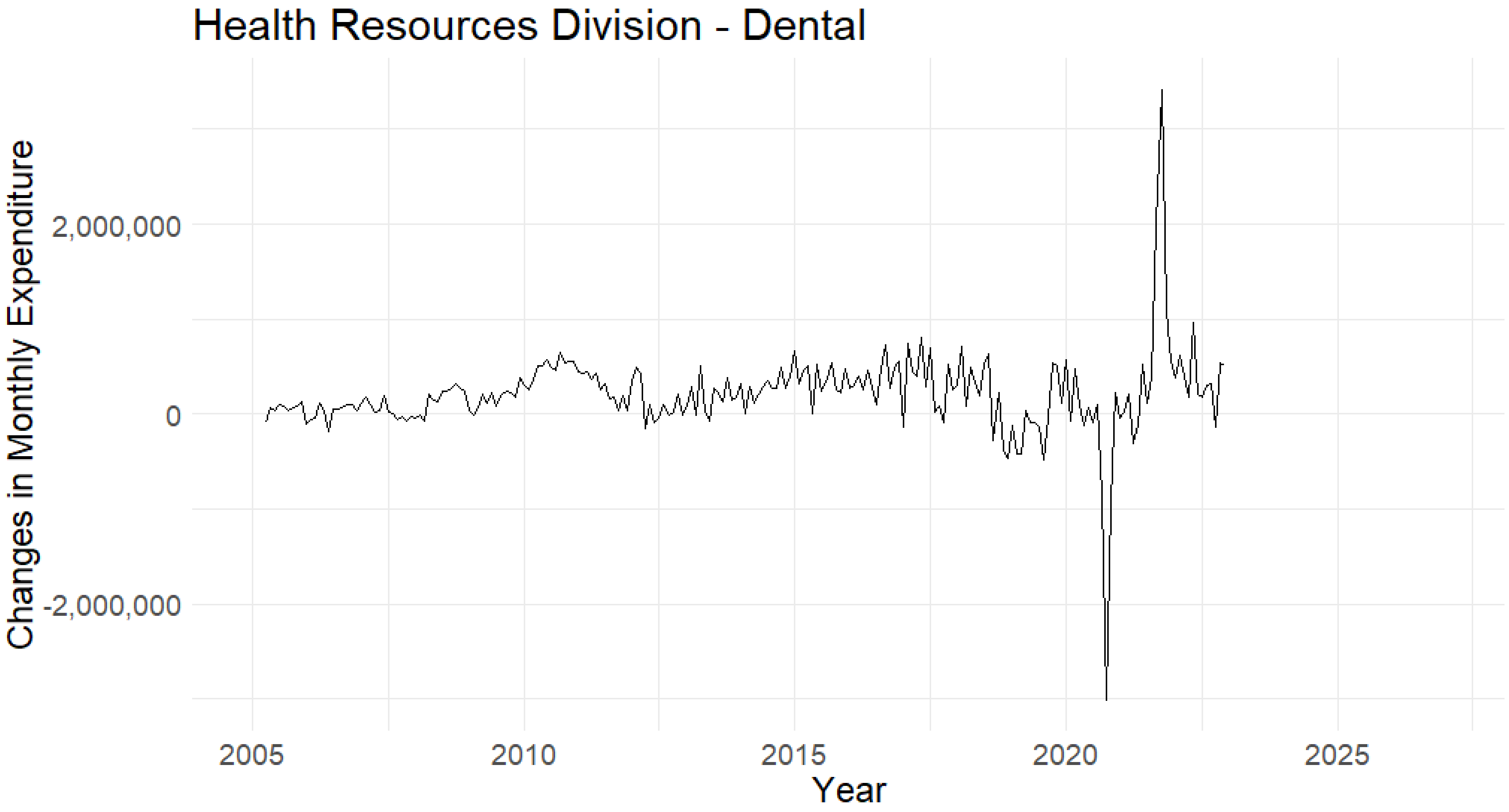
Appendix – Satisfying Stationarity

Seasonal Differencing.

A seasonal difference is the difference between an observation and the previous observation from the same season

Mathematically…

$$y'_t = y_t - y_{t-m}$$


Health Resources Division - Dental
Changes in Monthly Expenditure
2,000,000
0
-2,000,000
2005
2010
2015
2020
2025
Year

Appendix – ARIMA Statistical Form

$$y_t = \beta_1 * y_{t-1} + \alpha_1 * \varepsilon_{t-1} + \ldots + \beta_k * y_{t-k} + \alpha_k * \varepsilon_{t-k}$$

Appendix – Statistical Assumptions

Confidence Intervals assume that residuals are **normally distributed** and **uncorrelated**.

What does this mean?

- Normally Distributed: the error terms are distributed symmetrically
- Uncorrelated: there is no linear relationship between them

Mathematically…

1. $\varepsilon \sim N(\mu, \sigma^2)$
2. $cov(\varepsilon_t, \varepsilon_{t+k}) = E[\varepsilon_t \varepsilon_{t+k}] = 0$

Appendix – R Software

R is…

- Open-source software that can be extended via *packages*
- A language and environment for

  Data handling, storage and analysis

  Statistical Computation

  Looping, conditional, and user-defined recursive functions

  Graphical Display

- For more information visit the R-Project Website