Timeseries Analysis - the Tidyverse Way

George K Agyen

Table of contents

Welcome

A practical, beginner-friendly guide using tsibble, fable and friends

> " Target Audience: Beginners in R with basic knowledge of data frames and plotting, but little to no experience in time series analysis
>
> Tools Used: `tidyverse tsibble fable feasts lubridate`
>
> Datasets: Built-in data (`tsibble::tourism`, `datasets::AirPassengers`) and custom datasets (`local gh time series data in csv format`)

Have you ever looked at a line chart of sales, website traffic, stock prices or weather patterns and thought, *"I wonder what happens next?"* – then you are in the right place.

In this book, we will take your basic R skills and transform them into **real-world forecasting power** – all using the clean modern tools of the `Tidyverse`. No black-box algorithms or confusing jargon. Just a step by step journey into understanding patterns over time and predicting the future with confidence.

Whether you are a student, analyst or data enthusiast, time series skills are essential in your field – and surprisingly accessible. By the end of this guide, you will be able to:

- Visualise trends and seasonality like a pro

- Decompose complex patterns

- Build and compare forecasts using smart, automated models

- And most importantly explain your results clearly

So grab your pen and paper (whatever it is you want to grab), fire up RStudio and let's turn your curiosity into prediction.

1 Introduction

1.1 What is Time Series Data?

Time series data consists of observations recorded over time, usually at regular intervals (e.g., daily, monthly, yearly). Some examples include:

- Monthly rainfall totals
- Daily COVID-19 cases
- Hourly temperature readings
- Yearly population counts

What makes time series special is that time is not just a variable, It carries important structure and dependencies. What happened today can depend on what happened yesterday, last month or even last year. This temporal uniqueness is what makes time series very powerful.

1.2 Why Time Series Analysis Matters

Time series analysis helps us to **understand the past, monitor the present, and predict the future**. Some real world examples are:

- **Public Health:** Forecasting disease outbreaks or hospital admissions
- **Finance:** Predicting stock prices, currency exchange rates, or sales revenue.
- **Environmental Science:** Analysing temperature trends or rainfall patterns for climate studies
- **Engineering:** Monitoring sensor data to detect faults or changes in performance

In many cases, this analysis supports decision making — whether it is planning resources, anticipating risks, or detecting unusual patterns.

1.3 Traditional vs. Tidy Approach

Historically, time series in R has been handled using `ts` objects and packages like `forecast`. These are still useful, but they do not fit perfectly with the modern **tidy philosophy**; data frames, pipelines and consistent syntax.

We will follow the tidy time series workflow using the tidyverse ecosystem. The workflow uses packages from the **tidyverts** ecosystem:

- `tsibble`: A tidy data structure for time series (like a tibble but with special time handling)
- `fable`: For forecasting models (ARIMA, ETS, etc) in a tidy way.
- `feasts`: For exploratory analysis (seasonal plots, decomposition, autocorrelation).
- `lubridate`: For working with dates and times.
- `ggplot2`: For beautiful and flexible visualisations
- `dplyr/tidyr/tibble`: For general data wrangling

1.3.1 The Tidy Time Series Workflow

The tidyverse ecosystem in R emphasises clean, readable code and consistent data structure. For time series, the modern approach uses the `tidyverts` suite of packages.
The typical flow we will follow includes:

1. **Data Preparation:** Load and tidy data. Convert the data to a `tsibble` object so R knows how to handle time.
2. **Explore:** Visualise trends, seasonality, patterns
3. **Decompose:** Break down components (trends, seasonal, noise)
4. **Model:** Fit forecasting models (simple —> advanced)
5. **Forecast:** Generate future predictions
6. **Evaluate:** Check how good the forecasts are (accuracy assessment)

This workflow is clean, consistent and integrates smoothly with other tidyverse tools you might already know.

> 💡 Tip
>
> `tidyverse`: collection of packages designed for general data science
> `tidyverts`: collection of packages specifically for time series analysis
> They all follow the tidy philosophy, structure and grammar

1.3.2 Data We Will Use

We will start with some built-in datasets from the `tsibble` package so every on can follow along without downloading external files:

- `aus_production` — Quarterly production values for Australian industries.
- `tourism` — Quarterly overnight trips in different Australian regions.

Later, we shall also show how to use **real-life datasets** — for example, population growth or GDP growth data from Ghana, to make the examples relatable and practical

You can download the datasets used here `gh_data.csv` .

1.3.3 What You Will Need to Follow Along

1. **Basic R Knowledge:** You should know how to load packages, run functions, and work with data frames
2. **RStudio** installed for a smooth workflow
3. Internet connection (for package installation and possible data download)

1.3.3.1 By the end of this tutorial, you will be able to:

- Handle date/time data with ease
- Explore time series visually and statistically
- Build forecasts using tidyverse-style functions
- Apply your skills to your own datasets in research or work

Part I

Setting Up

Before diving into analysis, let;s make sure we have all the **tools** we need to get ready. We need to set up our R environment with the right tools. The `tidyverts` ecosystem — built around the `Tidyverse` — makes time series analysis intuitive, consistent and visual.
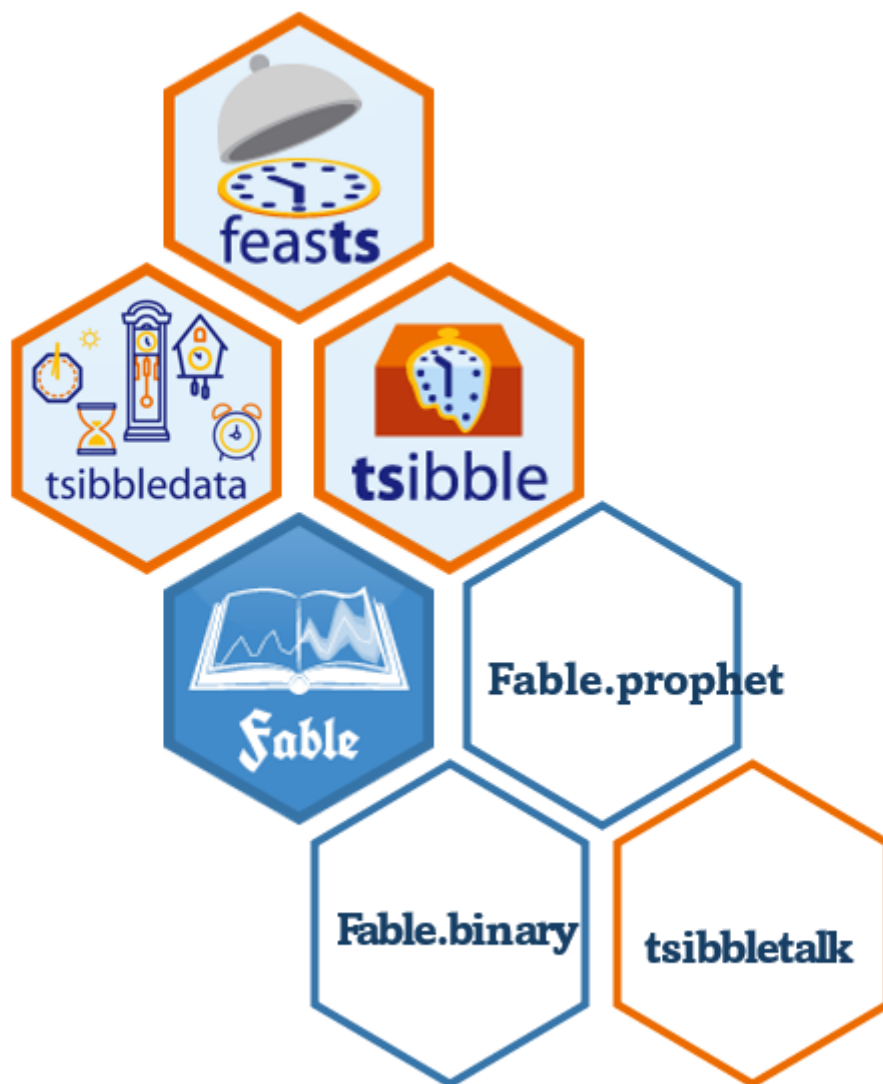
Figure 1.1: Tidyverts Package Collection

Figure 1.2: Tidyverse Package Collection

This section will cover installing and loading the packages we will use and checking that your R environment is prepared for a smooth time series work

We will address the following:

- Chapter **??**: This part will cover all the packages we will need, how to install and load them for use in our R environment.
- Chapter **??**: Here you will see how to keep things organised in a more seamless and orderly manner on your system.

After going through this section you should have been able to **install and load** all required packages, make RStudio r**eady for your time series coding** and then been able able to **successfully load a test** dataset.

2 Required Packages

As stated earlier we will be using a combination of `tidyverse` and time series specific packages (`tidyverts`)

Package	Purpose
tidyverse	Core suite for data wrangling and visualisation (`dplyr`, `tibble`, `ggplot2`, `readr`, etc)
tsibble	Tidy data frames for time series, handles dates, keys and indexing
fable	Forecasting models (ETS, ARIMA, Naive, etc) with tidy outputs
feasts	**F**eature **E**xtraction and **A**nalysis for **S**eries **T**ime **S**eries (plots, decomposition, autocorrelation, etc)
lubridate	Easy and readable date time manipulation
readr	Fast and easy data imports (csv, text data, etc)
readxl	Importing excel files (could be optional but very useful)

> **i** Note
>
> the `readr` and `lubridate` packages are already part of the **tidyverse package collections**. Once you install `tidyverse`, you do not need to install them again separately.

2.1 Installing the Package

3 Setting Up Your RStudio Environment

Part II

Working with Time Series in `tsibble`

4 Tidy Time Series Basics

5 Dealing with Time Gaps and Irregularities

6 Importing Data and Creating a tsibble