Module Code:	CS3DV
Assignment report Title:	Renewable Energy Adoption & Climate Change Response Visualization Report
Student Number (e.g. 25098635):	32011422
Actual hours spent for the assignment:	20h
Which AI tool used (if applicable):	

1. Introduction

This project explores how a raw dataset on renewable energy can be turned into a clear and meaningful visual analysis using Tableau. Although the data is initially provided as a simple flat file, it contains several important themes: Production levels, Investment patterns, Installed capacity, environmental impact, and employment figures – across different countries, energy sources, and years. The aim of this assignment is to organise this information in a way which makes it easier to explore and answer the two given questions: How production varies across countries and sources, and how these production trends evolve over time.

To achieve this, the report follows the typical workflow used in many real-world data visualisation tasks. The first step involves understanding the structure of the dataset and preparing it so that Tableau can use it effectively. After the work moves into dimensional modelling, where this file is reshaped conceptually into a star schema with clear dimensions and a central fact table. This structure forms the basis for creating two targeted data marts, small, focused subsets of the data designed to answer the two questions set out in the brief.

Once the data marts are in place, the focus shifts to visual design. Tableau is used to build separate views for each question, which are then combines into an interactive dashboard. The dashboard allows users to explore the data through filters and comparisons, making the patters easier to see. The report ends with a short analysis of the main trends, highlighting some of the differences between countries, the relative performance of each source, and how production has changed over the years.

Overall, the goal of this project is not only to produce a working dashboard but also to show a clear and logical process. From understanding the data, to modelling it, to designing visuals that help reveal the story behind the numbers.

2. Design Tasks

2.1 Design Star Schema

To support the analysis required for this coursework, I have designed a star scheme that reshapes the original flat CSV file into a structure which is easier to query, aggregate, and visualise. The dataset initially arrived as a single wide table containing country names, years, energy sources, policy description and multiple quantitative indicators related to renewable energy. As this format works for simple filtering, it however becomes inefficient when the goal is to slice the measures across different dimensions. A star scheme offers a cleaner and more analytical layout, so I reorganised the data around one central fact table and four surrounding dimensions.

The Fact Table (FactRenewableEnergy) holds all the numerical indicators that require aggregation or trend analysis. These measures change year-on-year and differ between countries and energy sources. I have included the following attributes as facts: Production(MWh), Investment(USD), Capacity(MW), Share(%), Created Jobs, CO2 Reduction. These values represent core performance metrics for renewable energy adoption, so grouping them together in the centre of the schema makes it easier to calculate totals, averages, or year-over-year-changes. The fact table also stores four foreign keys linking to the dimension tables: FK_Country, FK_Year, FK_Source, and FK_Policy.

For organizing the context around the measurements, I created four dimensions: DimCountry, DimDate, DimSource, DimPolicy. DimCountry contains the unique list of countries in the dataset. It serves as a stable lookup table, allowing the same country to be referenced across multiple years and energy sources. The primary key is simply the country name. DimDate uses yearly observations, so the date dimension is simple. I kept only the year field, which is sufficient to support trends and time-series analysis, the year acts as the primary key. DimSource stores the different renewable energy sources (Solar, Wind, Hydro, Geothermal and Biomass). By treating energy source as a dimension, I can easily compare performance across categories within a single visualisation or run deeper comparisons between technologies. DimPolicy contains the long raw dataset, with repeated policy descriptions. To make this more usable, I grouped these descriptions into broader policy categories (Solar policies, Wind policies, etc.). This dimension improves clarity and allows policy changes to be analysed without relying on messy text fields.

The main aim of the star schema is to simplify analytical queries. Tableau works best when the underlying model follows a fact-and-dimension structure, because it reduces duplication and avoids confusion caused by repeated text fields. Moving to this schema also makes the dashboards more intuitive: filters map directly to dimension tables, and charts pulling from the fact table became faster and cleaner.

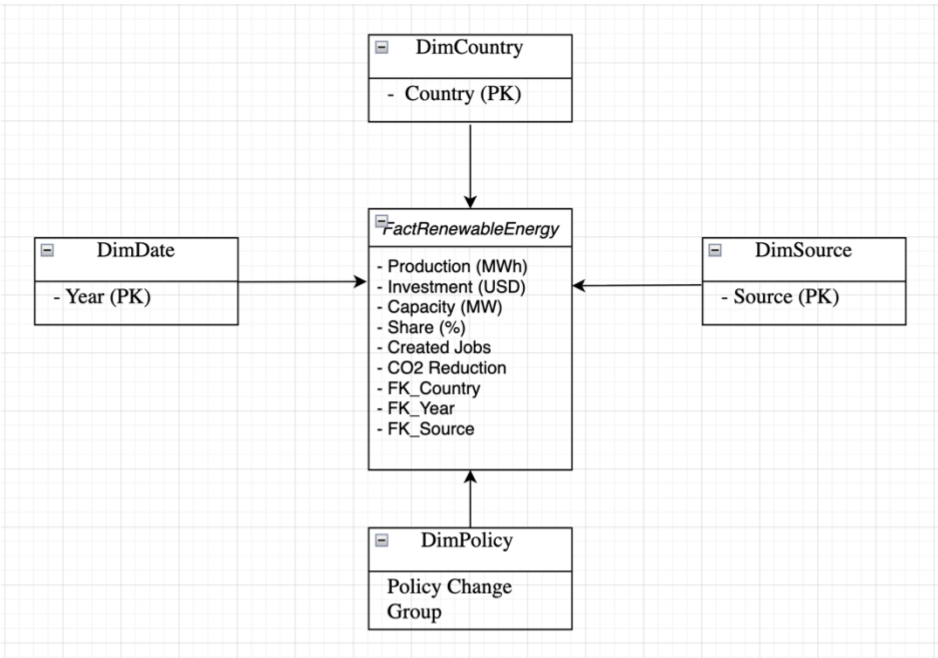

Figure 2.1.1: Star Schema

Finally, the star schema is shown in Figure 1. Which has visually illustrated the central fact table and its connections to the four dimensions using straightforward one-to-many relationships. This layout mirrors the structure used in most analytical systems and gives a clear, logical foundation for the visualisations that follow in the later sections.

2.2 Design a Datamart

I created a small data mart inside Tableau using the transformed star scheme from section 2.1. Although the original dataset was in a flat file format, it needed reconstructing and several cleaning steps before it could function reliably as a data mart (figure 2.2.1). The goal of this step is to ensure that each component of the scheme dimensions, facts, and hierarchies are both technically correct and analytically useful. The first issue I encountered was inconsistent formatting across several numeric fields. Investment values, installed capacity, and environmental indicators were initially stored as strings and contained commas or descriptive text. Tableau sometimes interprets these fields as text by defaults, which prevents aggregations and is why I was lead to this. To address this, I inspected each field and confirmed whether conversions were necessary. Fortunately, most fields loaded as cleanly as numeric values, but I still verified each of them through quick audit charts and summary checks. This process helped confirm that the dataset did not contain any outliers, missing years, or duplicated country source combinations.

Fields

Type	Field Name	Physical Table	Remote Field Name
	Country	renewable_energy_adoption.csv	Country/Region
#	Year	renewable_energy_adoption.csv	Year
Abc	Source	renewable_energy_adoption.csv	Energy Source
#	Production(MWh)	renewable_energy_adoption.csv	Energy Production (MWh)
#	Investment(USD)	renewable_energy_adoption.csv	Investment in Renewable Infrastructure (USD)
Abc	Policy Changes	renewable_energy_adoption.csv	Policy Changes
#	CO2 Reduction	renewable_energy_adoption.csv	Environmental Impact (CO2 Reduction)
#	Capacity(MW)	renewable_energy_adoption.csv	Installed Capacity (MW)
#	Share(%)	renewable_energy_adoption.csv	Renewable Energy Share (%)
#	Created Jobs	renewable_energy_adoption.csv	Jobs Created
	Policy Changes (group)	Group	Policy Changes (group)

Figure 2.2.1: Cleaned numeric fields after verifying and normalizing data types in Tableau

The dataset included long descriptive policy-change labels for every observation. These repeated sentences made filtering difficult and produced cluttered axes when graphed. To create a more coherent data mart, I grouped these policy descriptions into broader categories such as Solar Policy Change, Wind Policy Change, etc (Figure 2.2.2). This grouping now serves as the DimPolicy dimension and drastically simplified later in the dashboard interactions. Once the fields were cleaned, I extracted the unique values for each dimension and checked for uniqueness. I ensured that the fact table referenced only valid keys. This validation step was important as it avoided mismatches or broken filters later in Tableau.

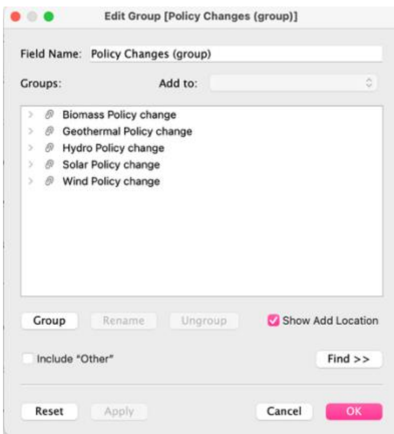

Figure 2.2.2 Grouping of raw policy descriptions into high-level policy categories

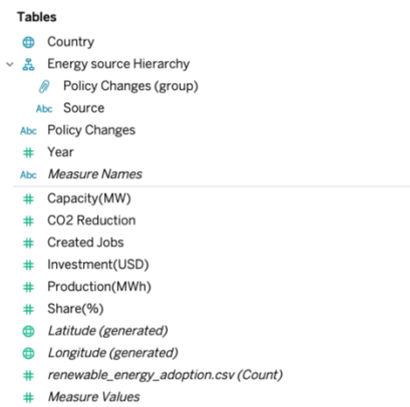

Figure 2.2.3 Dimension tables extracted from the original dataset and structures for analytical use

The Fact Table was constructed by retaining only the quantitative measures and the foreign-key fields linking back to the four dimensions. Because Tableau does not physically materialize tables in the same way as SQL databases, the “fact table” exists logically through the measures and their relationships to the dimensions. To validate the correctness of this structure, I created temporary exploratory sheets that plotted each measure against its corresponding dimensions. These sheets acted as mini data-quality checks. To confirm the data mart worked as intended I created a series of audit charts: Audit_Production (Figure 2.2.4), Audit_Investement (Figure 2.2.5), Audit_Capacity (Figure2.2.6). each chart was coloured by energy source to reveal whether the relationships were consistent across all groups. This helped to confirm that no missing joins, broken hierarchies, or typos.

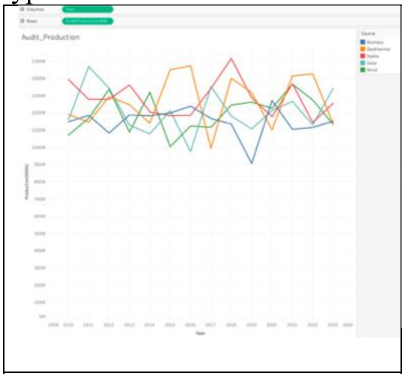		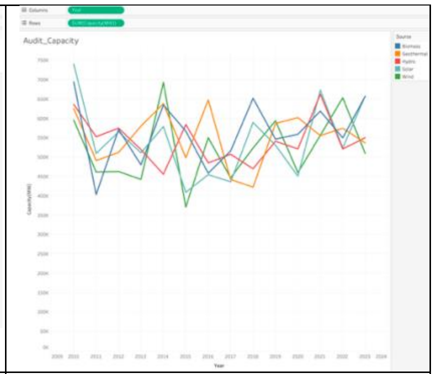
Figure 2.2.4: Audit chart confirming consistent production values across all sources.	*Figure 2.2.5: Audit Investment*	*Figure 2.2.6: Audit Capacity*

After the cleaning, grouping, and structural validation, the dataset functioned as a proper analytical data mart. Each filter in Tableau corresponded cleanly to one of the dimensions, and every visualization could use the fact table measures without errors or ambiguous field names. With the data mart in place, I was able to proceed to dashboard development with confidence that the underlying model was stable and analytically sound.

2.3 Design a Dashboard

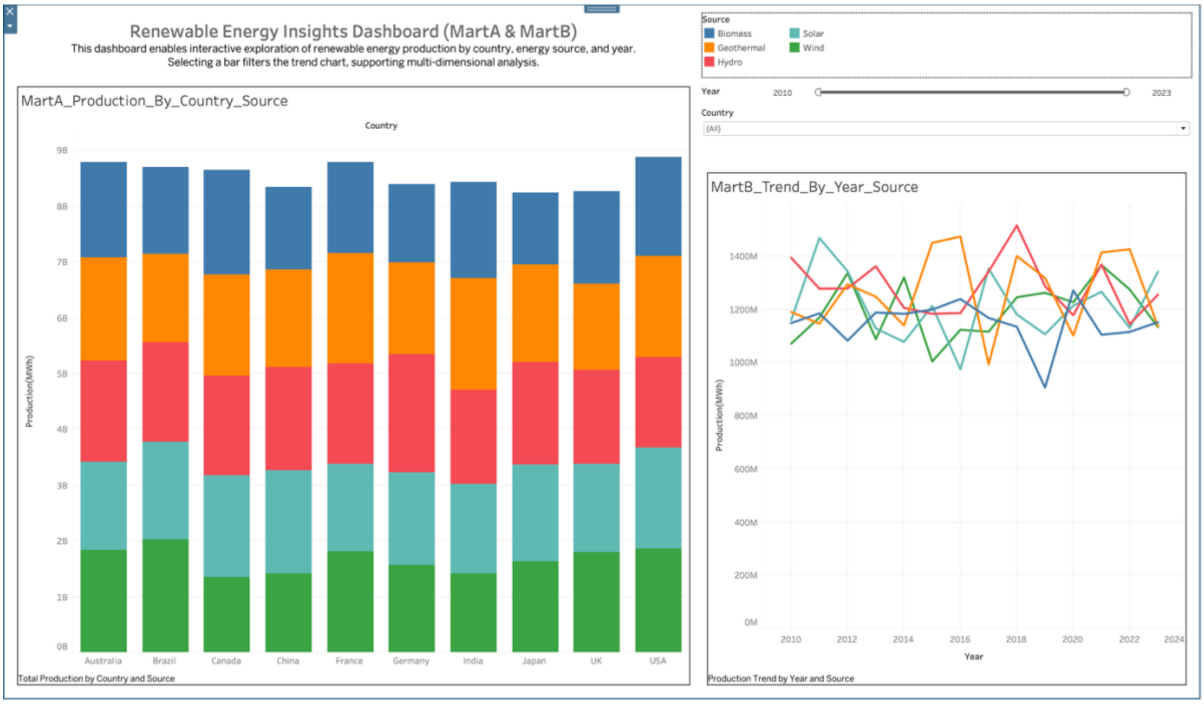

Figure 2.3.1: Dashboard

The purpose of this dashboard was to create an interactive and intuitive interface that allows users to explore renewable energy patterns across countries, sources, and years. The final design integrates both MartA and MartB outputs into a single analytical view, enabling cross-filtering and drill-down behavior.

The dashboard is divided into two main visual components: MartA Production by Country and Source and MartB Trend by year and source (Figure 2.3.1). MartA is a stacked bar chart summarizing total energy production for each country, broken down by energy source. This chart allows quick comparison of national energy profiles and relative reliance on solar, wind, hydro, geothermal, and biomass. MartB is a multiline trend chart that visualizes how production changes over time for each energy source. This supports temporal analysis and identifies long-term growth or decline with each energy category. These two charts are placed side to side, providing both a geographic and temporal perspective in a single dashboard.

I implemented several interactive elements, such as filters for source, year and country that were placed at the top of the dashboard to let users quickly adjust the scope of the analysis. A highlight action was configured so that selecting a bar segment in Mart A automatically highlights the corresponding trend in Mart B. this linkage allows users to drill down into specific country-source combinations. Another implementation I added were custom tooltips on both charts to provide additional context, including detailed values such as production(MWh), capacity(MW), investment(USD), and CO2 reduction. The dashboard uses responsive layout behavior, ensuring the charts resize properly and remain readable across different screen sizes.

The design focuses on clarity, usability, and minimal cognitive load. A vertical layout was chosen for MartA to maximise visibility of country clutter. Colour coding is consistent across both charts (same palette for energy sources), which supports quick pattern recognition. Therefore, this dashboard provides a cohesive and interactive visual environment that allows users to explore renewable energy data form multiple angles and gain actionable insights.

3. Conclusion and Recommendations

The analysis from this dashboard confirmed that renewable-energy performance varies significantly by both country and source, with some nations demonstrating consistently higher production levels across all technologies. The dashboard made these differences clear by combining a country-based production view with a yearly trend analysis, allowing patterns such as growth cycles, volatility, and relative contribution of each source to become visible. However a small number of fields contained missing or inconsistent values, the marts and star scheme ensured that the dataset remained structures, reliable, and easy to analyse. To strengthen this work further I would firstly expand the data to include additional indicators such as cost per unit, land use, or policy intensity, which would allow more room for a deeper understanding of why performances differs across countries. Secondly the missing values, particularly in the CO2 reduction field where it should've been cleaned or imputed to support more complete environmental-impact insights. Overall, this project delivers a clear and usable analytical view of global renewable-energy trends, while also highlighting opportunities for richer future analysis and the dashboards interactive features allow users to explore country-source relationships dynamically, adding analytical value beyond static charts.

4. Reflection on Learning Experience

Working throughout this coursework gave me a much clearer understanding of how data modelling and visual analytics fit together in practice. Before starting I mainly viewed tableau as a chart building platform, but this project made me realise the importance of shaping the data structure first. Creating the star schema and marts forced me to think carefully about what should be treated as a dimension versus a fact, and I realise how much easier analysis becomes once the model mirrors the way the data will be queried.

References

Slalom (2025) 'Optimizing data models for Tableau dashboard performance', Medium*. Available at: https://medium.com/slalom-blog/optimizing-data-models-for-tableau-dashboard-performance-fdbd71a33a0c*

University of Reading (2024). CS3DV Lecture Slides: Dimensional Modelling and Data Mart Schema*. Available in module materials.*

Verissimo, B. (2025) 'Data Modeling in Tableau: A Practical Guide for BI Engineers', Medium*. Available at: https://medium.com/@bruno.verissimo/data-modeling-in-tableau-a-practical-guide-for-bi-engineers-a13dab90d347*

Gitlab link: https://csgitlab.reading.ac.uk/rp011422/cs3dv_cw_pranavi_rawal/-/tree/main