Clean-Tech Patent Analysis

2023-07-25

EDA: What Clean Tech Sector is Growing the Most?

In this analysis, I aim to illustrate my capabilities in data cleaning and preliminary analysis using a dataset containing patents registerd with OECD under the Y02 scheme developed by the European Patent Office for technical attributes of technologies that are related to climate change mitigation.

As an aspiring data analyst with a keen interest in the environment and sustainable solutions, I chose to work with a dataset on patents related to clean technology. This choice not only allowed me to contribute to a topic of personal interest but also to answer vital business questions relevant to startups and corporations alike in the clean tech industry.

Through this project, I explore several critical business questions: Which countries are leading in clean tech patents as of 2021? What are the prevalent trends in patenting over time? Which sectors within clean technology are the most innovative? All these questions, and more, can provide essential insights to shape business strategies, particularly for startups looking to carve out a niche in the competitive landscape.

This analysis, thus, is not just a demonstration of my data wrangling and exploratory skills, but it also underscores the power of data in answering crucial business questions and making informed decisions in an industry poised for substantial growth in the coming years.

The data used for this analysis was collected from the Energy Technology Patents Data Explorer. The column labled "value" represents the number of patents registered under a specific designation, country, and year.

Cleaning the Data

In this project, I focused on data cleaning and preliminary analysis using an international patent dataset. My goal was to prepare this data for future in-depth analysis, while also showcasing my skills in managing and handling complex, real-world datasets.

I started by importing a dataset containing information about patents from various countries and technology sectors. The data was read into R, and I began the initial exploration by viewing the first few rows of the dataframe with the head() function. This gave me a quick snapshot of the data structure and what kind of information was available.

One crucial step in data cleaning was handling missing values. I used colSums(is.na(patents_df)) to find columns with missing values. Then, I filtered out rows where 'countryISO' was 'WORLD' or 'techSector' was 'Total - all technologies'. The reason for this was twofold. First, 'WORLD' isn't a specific country and could distort country-specific analyses. Second, I aimed to analyze specific technology sectors, making 'Total - all technologies' data unnecessary. I could always go back and refer to these variables if I needed to.

After removing these rows, I checked for missing values again. To handle them, I filled missing values in the 'indicator' and 'category' columns with 'Unknown'. This decision allowed me to preserve these rows' other information while clearly marking where data was missing.

Duplicate entries in a dataset can also cause skewed analysis results. So, I used the anyDuplicated() function to check if any such entries were present in the dataframe.

To verify that the data types of each column were as expected (for example, numeric, character, etc.), I utilized the str() function. This check is important because improper data types can lead to errors during analysis or produce incorrect results.

Finally, after all these cleaning steps, I printed the first few rows of the cleaned dataset. This provided a quick overview of what the cleaned data looked like and confirmed the successful execution of the cleaning procedures.

The clean and prepared data is now ready for further exploratory data analysis and modeling. In future steps, I will leverage this data to answer critical business questions, providing valuable insights into the clean technology industry's patent landscape.

Preliminary Analysis

In the Preliminary Analysis phase of this project, I conducted a series of operations aimed at gaining a deeper understanding of the dataset's structure, identifying key variables, and exposing important trends within the data.

First, I listed all unique technology sectors present in the dataset using the unique() function. This step allowed me to comprehend the data's diversity and plan which sectors might be interesting to study further, especially in relation to clean technologies.

To better understand the distribution of patents, I created a histogram and a boxplot of the 'value' variable, which represents the number of patents. These visualizations provided insight into the data's spread, central tendency, and potential outliers.

Following this, I created a boxplot of patent values grouped by countries. This visualization helped to explore the variation in patent values across different countries, allowing for a comparative perspective on technological advancement across geographical regions.

Aiming to provide insights for the key business questions, I focused on summarizing the 'Total - all technologies' sector, excluding the 'WORLD' country code. The bar plot created from this data gives a visualization of the total number of patents per country, showcasing where the majority of patenting activity has taken place.

To specifically address the focus on clean energy technologies, I filtered and grouped the data to display the top ten countries with the most clean energy patents. By visualizing this data in a bar plot, it clearly showed which countries are leading in clean energy patents, potentially indicating where the most significant advancements in clean technologies are taking place.

For a more detailed analysis, I arranged the cleaned data in decreasing order of the number of patents, giving a view of the most patented technologies irrespective of the year.

In summary, this preliminary analysis was fundamental to understand the structure and key trends within the patent data, and to address the project's core business questions around clean technology innovation.

Trends

In this section of the analysis, I wanted to understand the trends in patents over time across different technology sectors on a global scale. This understanding could provide insights into the progression of innovation within these sectors. I iterated over each unique technology field in the dataset. For each technology field, I filtered the data for only the instances where the technology sector matched the field in the current iteration and the country was marked as "WORLD", representing the global data.

Next, I checked if the number of observations for the particular technology field was more than one, ensuring there was sufficient data for analysis. If this condition was met, I grouped the filtered data by 'year' and calculated the total number of patents for each year. I then created a line graph to visualize the trend of patents over the years for each technology field.

To further assist with interpreting the trends, I added a regression line to the plots using the geom_smooth function. This line, generated using a linear model (method = "lm"), provided an understanding of the overall trend in the patent data for each technology sector over time. This could help identify sectors that are seeing growth or decline in terms of innovation on a global scale.

The red regression line adds a clear visual representation of the trends and helps answer some of our initial business questions regarding the progress and change in different technology sectors. Each graph was titled appropriately for clarity.

This section of the analysis helped me understand the patent trends in the different technology sectors and identify potential areas of growth and interest. It also served as a way to visually present the data in a way that would be easy for stakeholders to understand.

Conclusion

In this project, I've performed a comprehensive analysis of clean technology patents data provided by the OECD and EIA. Through data cleaning, preliminary analysis, and visualization, I've extracted valuable insights about the trends in patents over years in various tech sectors globally. The analysis showcases my skills in data cleaning, analysis, and visualization using R and provides answers to crucial business questions posed at the beginning of the project.

The trends observed from the analysis indicate a significant interest and growth in many of the clean technology sectors worldwide. The data suggests that innovations in the clean tech sectors have been increasing, which could be due to growing environmental awareness, international policies, and worldwide collaboration.

For a startup or corporation interested in the clean tech sector, the insights derived from this project could provide valuable guidance for where to enter the market. For example there is positive trends in patents registered under Grid, Agricultural Energy Effeciency, and E-Mobility. They could consider focusing on the tech sectors showing the most growth, as indicated by the increasing number of patents. Moreover, the United States, Japan, China, and Germany are leading in terms of the number of patents, suggesting a more active and supportive environment for clean tech innovations and funding.

Next Steps

While the analysis has provided insightful information, there are several directions that future work could take:

Drill Down Analysis: Perform a more in-depth analysis at a country level. This could involve analyzing the trends in different tech sectors within a specific country or comparing the growth between different countries.

Predictive Analysis: Use the trends observed in the data to forecast future growth in the number of patents in different tech sectors. This could help organizations to strategize their innovation efforts.

Explore Correlations: Look for potential correlations between the growth of patents in a tech sector and other economic or environmental indicators. This could provide additional insights into factors influencing growth in these sectors.

Policy Analysis: Evaluate the effect of specific policies on the growth of patents in clean tech. This could help identify favorable conditions for innovation in clean technology.

These future steps could further improve the depth and breadth of the analysis and provide even more valuable insights for startups or corporations interested in the clean tech sector.

Code:

```
## Clean Data

# Display the first few rows of the dataframe
head(patents_df)
```

```
## # A tibble: 6 x 10
##   country        countryISO techSector      topic indicator category unit   year
##   <chr>          <chr>      <chr>           <chr> <chr>     <chr>    <chr> <dbl>
## 1 Netherlands    NLD        Hydrogen fuel ~ Pate~ Detail    Process~ Numb~  2000
## 2 United States  USA        Nuclear         Pate~ Detail    Clean E~ Numb~  2000
## 3 France         FRA        Nuclear         Pate~ Detail    Clean E~ Numb~  2000
## 4 Japan          JPN        Grid            Pate~ Detail    Clean E~ Numb~  2000
## 5 Australia      AUS        Renewable ener~ Pate~ Detail    Clean E~ Numb~  2000
## 6 Czech Republic CZE        Air - rail - m~ Pate~ Detail    Clean E~ Numb~  2000
## # i 2 more variables: observations <dbl>, value <dbl>
```

```
patents_df
```

```
## # A tibble: 18,881 x 10
##    country        countryISO techSector     topic indicator category unit   year
##    <chr>          <chr>      <chr>          <chr> <chr>     <chr>    <chr> <dbl>
##  1 Netherlands    NLD        Hydrogen fuel~ Pate~ Detail    Process~ Numb~  2000
##  2 United States  USA        Nuclear        Pate~ Detail    Clean E~ Numb~  2000
##  3 France         FRA        Nuclear        Pate~ Detail    Clean E~ Numb~  2000
##  4 Japan          JPN        Grid           Pate~ Detail    Clean E~ Numb~  2000
##  5 Australia      AUS        Renewable ene~ Pate~ Detail    Clean E~ Numb~  2000
##  6 Czech Republic CZE        Air - rail - ~ Pate~ Detail    Clean E~ Numb~  2000
##  7 Korea          KOR        Agriculture e~ Pate~ Detail    Clean E~ Numb~  2000
##  8 United States  USA        Coal-to-liqui~ Pate~ Detail    Process~ Numb~  2000
##  9 United States  USA        Gaseous fuel ~ Pate~ Detail    Transmi~ Numb~  2000
## 10 Belgium        BEL        Conventional ~ Pate~ Detail    Upstream Numb~  2000
## # i 18,871 more rows
## # i 2 more variables: observations <dbl>, value <dbl>
```

```
# Check for missing values
colSums(is.na(patents_df))
```

```
##      country   countryISO   techSector        topic    indicator     category
##          928            0            0            0          770          770
##         unit         year observations        value
##            0            0            0            0
```

```r
# Remove rows where countryISO is 'WORLD' or  techSector is 'Total - all technologies'
patents_filtered <- patents_df[!(patents_df$countryISO == 'WORLD' | patents_df$techSector == 'Total - al

# Print the filtered dataset
patents_filtered
```

```
## # A tibble: 17,205 x 10
##    country        countryISO techSector     topic indicator category unit   year
##    <chr>          <chr>      <chr>          <chr> <chr>     <chr>    <chr> <dbl>
##  1 Netherlands    NLD        Hydrogen fuel~ Pate~ Detail    Process~ Numb~  2000
##  2 United States  USA        Nuclear        Pate~ Detail    Clean E~ Numb~  2000
##  3 France         FRA        Nuclear        Pate~ Detail    Clean E~ Numb~  2000
##  4 Japan          JPN        Grid           Pate~ Detail    Clean E~ Numb~  2000
##  5 Australia      AUS        Renewable ene~ Pate~ Detail    Clean E~ Numb~  2000
##  6 Czech Republic CZE        Air - rail - ~ Pate~ Detail    Clean E~ Numb~  2000
##  7 Korea          KOR        Agriculture e~ Pate~ Detail    Clean E~ Numb~  2000
##  8 United States  USA        Coal-to-liqui~ Pate~ Detail    Process~ Numb~  2000
##  9 United States  USA        Gaseous fuel ~ Pate~ Detail    Transmi~ Numb~  2000
## 10 Belgium        BEL        Conventional ~ Pate~ Detail    Upstream Numb~  2000
## # i 17,195 more rows
## # i 2 more variables: observations <dbl>, value <dbl>
```

```r
# Check for missing values
colSums(is.na(patents_filtered))
```

```
##      country   countryISO   techSector        topic    indicator     category
##            0            0            0            0            0            0
##         unit         year observations        value
##            0            0            0            0
```

```r
# Check for duplicates
anyDuplicated(patents_filtered)
```

```
## [1] 0
```

```r
# Fill missing values in the 'indicator' and 'category' columns with 'Unknown'
patents_filtered$indicator[is.na(patents_filtered$indicator)] <- 'Unknown'
patents_filtered$category[is.na(patents_filtered$category)] <- 'Unknown'

# Check data types of the columns
str(patents_filtered)
```

```
## tibble [17,205 x 10] (S3: tbl_df/tbl/data.frame)
##  $ country     : chr [1:17205] "Netherlands" "United States" "France" "Japan" ...
##  $ countryISO  : chr [1:17205] "NLD" "USA" "FRA" "JPN" ...
##  $ techSector  : chr [1:17205] "Hydrogen fuel production" "Nuclear" "Nuclear" "Grid" ...
##  $ topic       : chr [1:17205] "Patents" "Patents" "Patents" "Patents" ...
##  $ indicator   : chr [1:17205] "Detail" "Detail" "Detail" "Detail" ...
##  $ category    : chr [1:17205] "Processing downstream" "Clean Energy" "Clean Energy" "Clean Energy"
##  $ unit        : chr [1:17205] "Number" "Number" "Number" "Number" ...
##  $ year        : num [1:17205] 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 ...
##  $ observations: num [1:17205] 0 0 0 0 0 0 0 0 0 0 ...
##  $ value       : num [1:17205] 2 43.5 26 12.5 3 ...
```

```r
# Display the cleaned dataframe
head(patents_filtered)
```

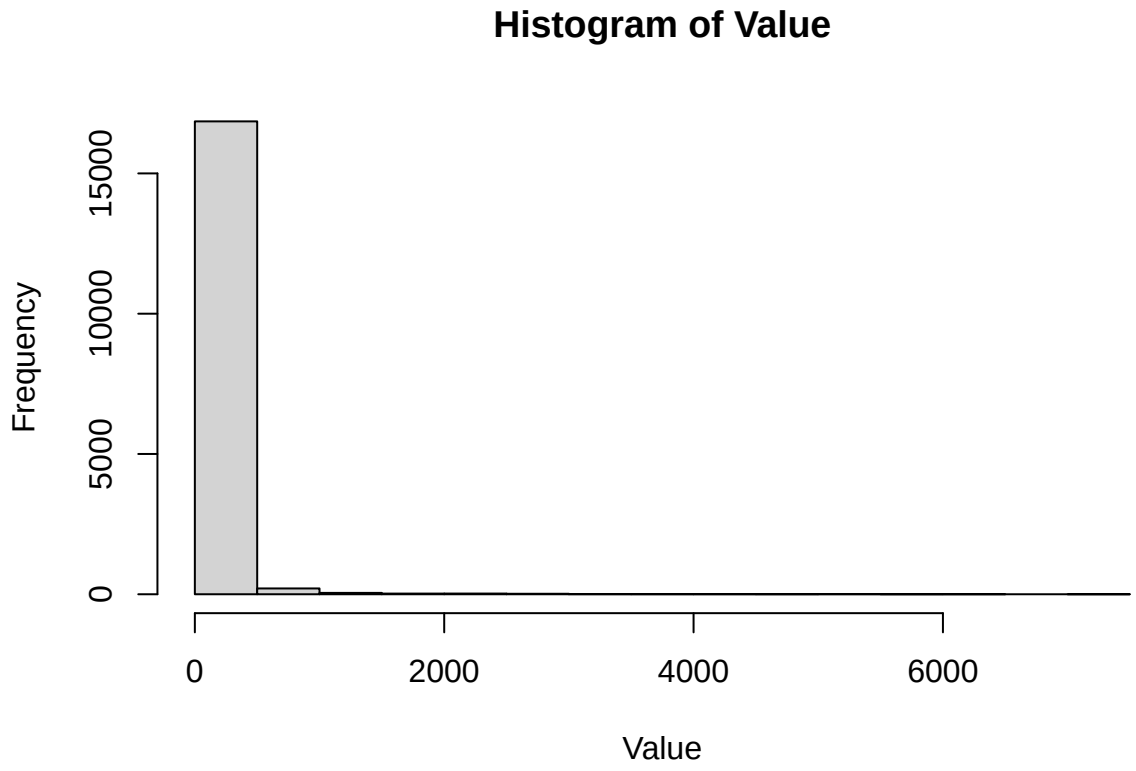
```
## # A tibble: 6 x 10
##   country        countryISO techSector      topic indicator category unit   year
##   <chr>          <chr>      <chr>           <chr> <chr>     <chr>    <chr> <dbl>
## 1 Netherlands    NLD        Hydrogen fuel ~ Pate~ Detail    Process~ Numb~  2000
## 2 United States  USA        Nuclear         Pate~ Detail    Clean E~ Numb~  2000
## 3 France         FRA        Nuclear         Pate~ Detail    Clean E~ Numb~  2000
## 4 Japan          JPN        Grid            Pate~ Detail    Clean E~ Numb~  2000
## 5 Australia      AUS        Renewable ener~ Pate~ Detail    Clean E~ Numb~  2000
## 6 Czech Republic CZE        Air - rail - m~ Pate~ Detail    Clean E~ Numb~  2000
## # i 2 more variables: observations <dbl>, value <dbl>
```

```r
# List unique sectors in the dataframe
tech_fields <- unique(patents_filtered$techSector)
print(tech_fields)
```

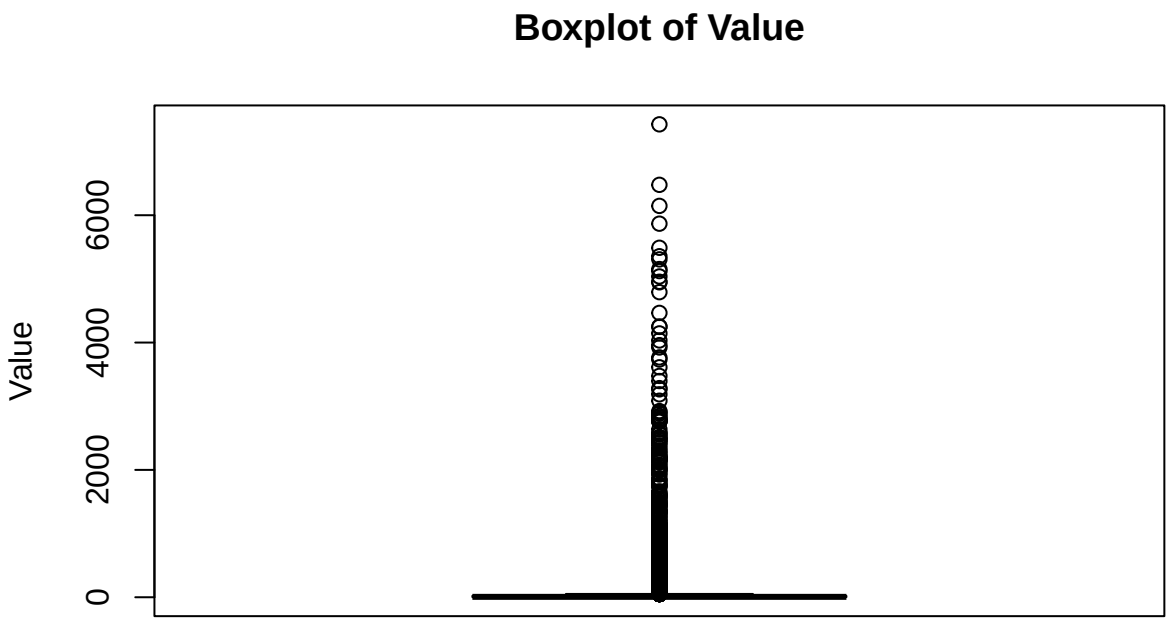
```
##  [1] "Hydrogen fuel production"
##  [2] "Nuclear"
##  [3] "Grid"
##  [4] "Renewable energy integration in buildings"
##  [5] "Air - rail - marine"
##  [6] "Agriculture energy efficiency"
##  [7] "Coal-to-liquids and gas-to-liquids"
##  [8] "Gaseous fuel distribution"
##  [9] "Conventional oil and gas exploration and extraction"
## [10] "Processing and downstream"
## [11] "Oil refining"
## [12] "Underground gaseous fuel storage"
## [13] "Industry energy efficiency or substitution"
## [14] "Bioenergy"
## [15] "Building energy efficiency"
## [16] "Gas conditioning"
## [17] "Solar"
## [18] "Coal-to-gas"
## [19] "Wind"
## [20] "Hydrogen and fuel cells"
## [21] "Transmission and distribution"
## [22] "Upstream"
## [23] "Liquid fuel distribution (gas stations)"
## [24] "Clean energy patents"
## [25] "Unconventional oil and gas exploration and extraction"
## [26] "Storage (not e-mobility)"
## [27] "Vehicle fuel efficiency"
## [28] "Fossil fuel patents"
## [29] "Stationary tank storage for gases"
## [30] "Other renewables"
## [31] "Renewables"
## [32] "e-Mobility"
## [33] "Liquified gaseous fuel shipping"
## [34] "Energy efficiency"
## [35] "Coal and solid fuels exploration and mining"
```

```
## [36] "Carbon capture and storage"
## [37] "Compressed gaseous fuel shipping"
## [38] "Solid fuel conditioning"
## [39] "Liquid fuel pipelines"
## [40] "Gas fuel pipelines"
## [41] "Rail tanker liquid fuels transport"
## [42] "Road tanker gaseous fuels transport"
## [43] "Underground liquid fuels storage"
## [44] "Road tanker liquid fuels transport"
## [45] "Liquid fuel tanker shipping"
## [46] "Rail solid fuels transport"
## [47] "Stationary tank storage for liquids"
## [48] "Solid fuel shipping"
## [49] "Rail tanker gaseous fuels transport"
```

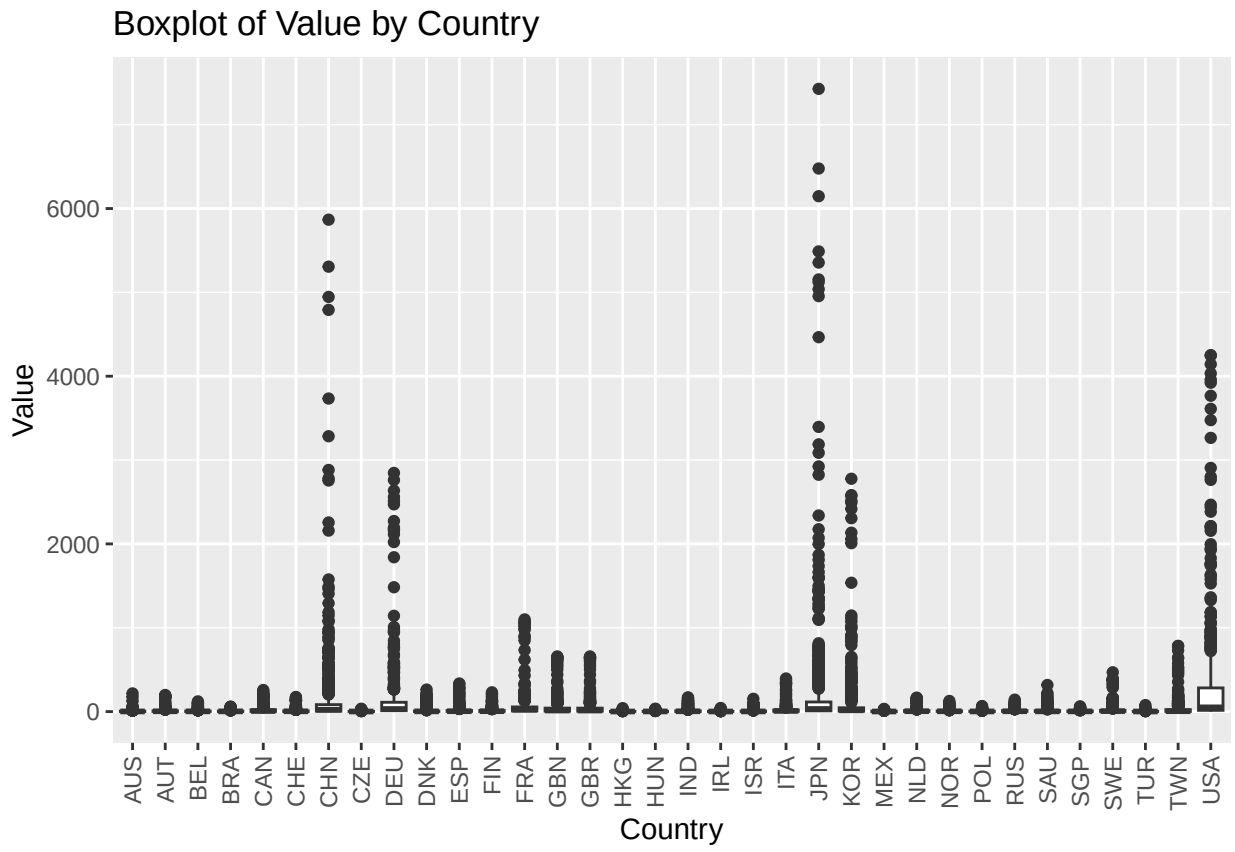
```r
# Display Histogram of Patent Data
hist(patents_filtered$value, main = "Histogram of Value", xlab = "Value")
```




```r
# Display Boxplot of Patent Data
boxplot(patents_filtered$value, main = "Boxplot of Value", ylab = "Value")
```




```
ggplot(patents_filtered, aes(x = countryISO, y = value)) +
  geom_boxplot() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1)) +
  labs(title = "Boxplot of Value by Country", x = "Country", y = "Value")
```

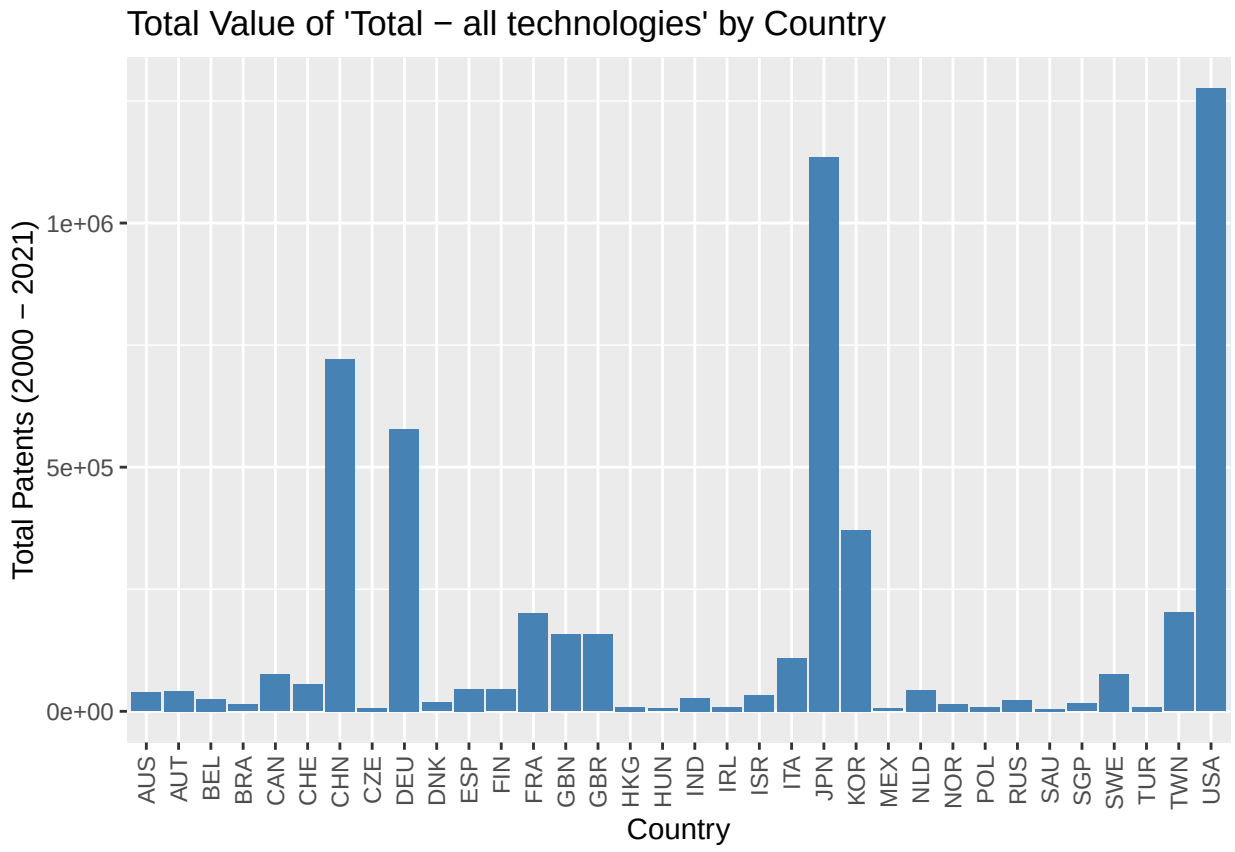



```
# arrange rows in order of decreasing values in the patents column while ignoring the year
top_patents <- arrange(patents_filtered, desc(value))
head(top_patents, 100)
```

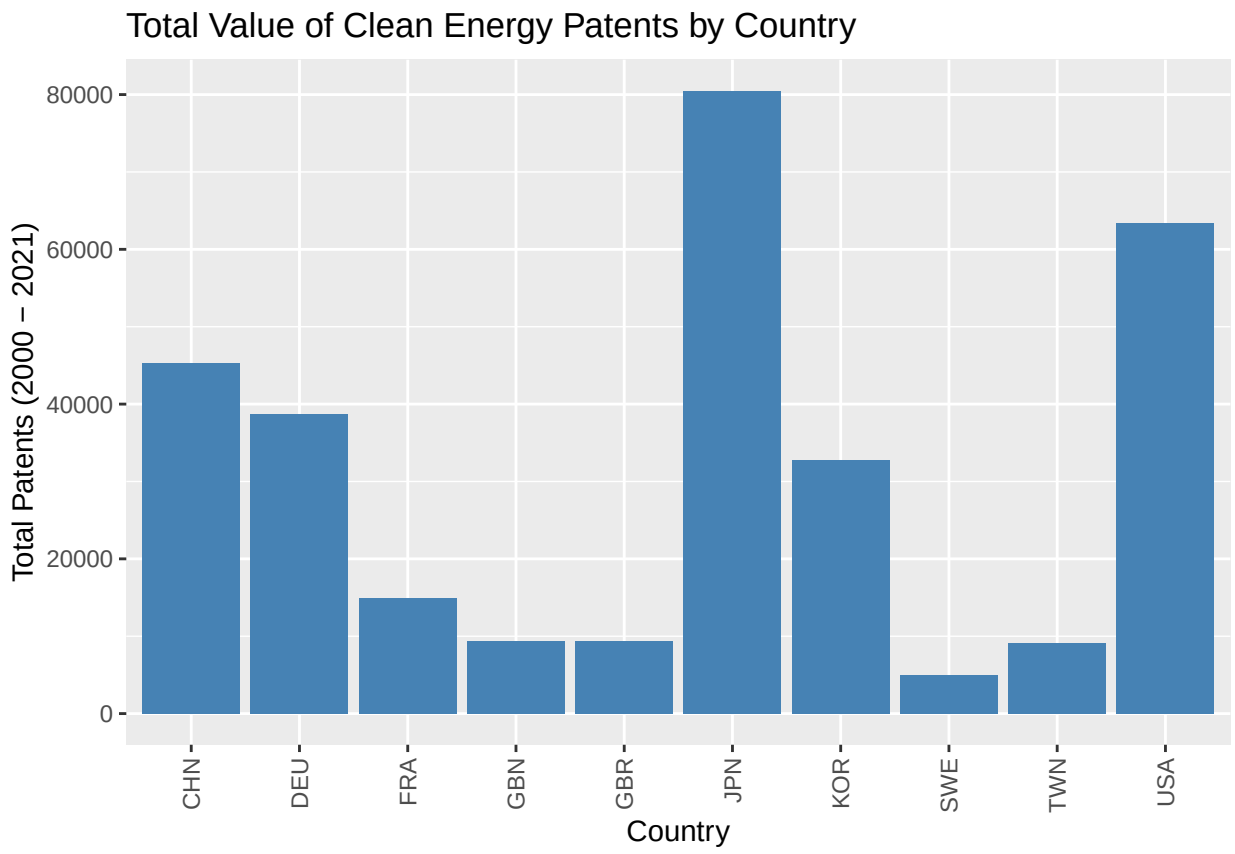
```
## # A tibble: 100 x 10
##    country            countryISO techSector topic indicator category unit   year
##    <chr>              <chr>      <chr>      <chr> <chr>     <chr>    <chr> <dbl>
##  1 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2011
##  2 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2012
##  3 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2010
##  4 People's Republic~ CHN        Clean ene~ Pate~ Total     Total    Numb~  2018
##  5 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2013
##  6 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2017
##  7 People's Republic~ CHN        Clean ene~ Pate~ Total     Total    Numb~  2017
##  8 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2015
##  9 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2016
## 10 Japan              JPN        Clean ene~ Pate~ Total     Total    Numb~  2018
## # i 90 more rows
## # i 2 more variables: observations <dbl>, value <dbl>
```

```
# Filter the data for 'Total - all technologies' and calculate total values per country
total_tech <- patents_df %>%
  filter(techSector == 'Total - all technologies', countryISO != 'WORLD') %>%
  group_by(countryISO) %>%
  summarize(total_value = sum(value, na.rm = TRUE))
```

```r
# Bar plot of total value per country
ggplot(total_tech, aes(x = countryISO, y = total_value)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  labs(title = "Total Value of 'Total - all technologies' by Country", x = "Country", y = "Total Patents
```




```r
# Rank the top ten countries with the most patents recorded in clean energy
clean_energy <- patents_df %>%
  filter(techSector == 'Clean energy patents', countryISO != 'WORLD') %>%
  group_by(countryISO) %>%
  summarize(total_value = sum(value, na.rm = TRUE)) %>%
  arrange(desc(total_value)) %>%
  slice_head(n = 10)

# Bar plot of top ten clean energy patents
ggplot(clean_energy, aes(x = countryISO, y = total_value)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) +
  labs(title = "Total Value of Clean Energy Patents by Country", x = "Country", y = "Total Patents (200
```



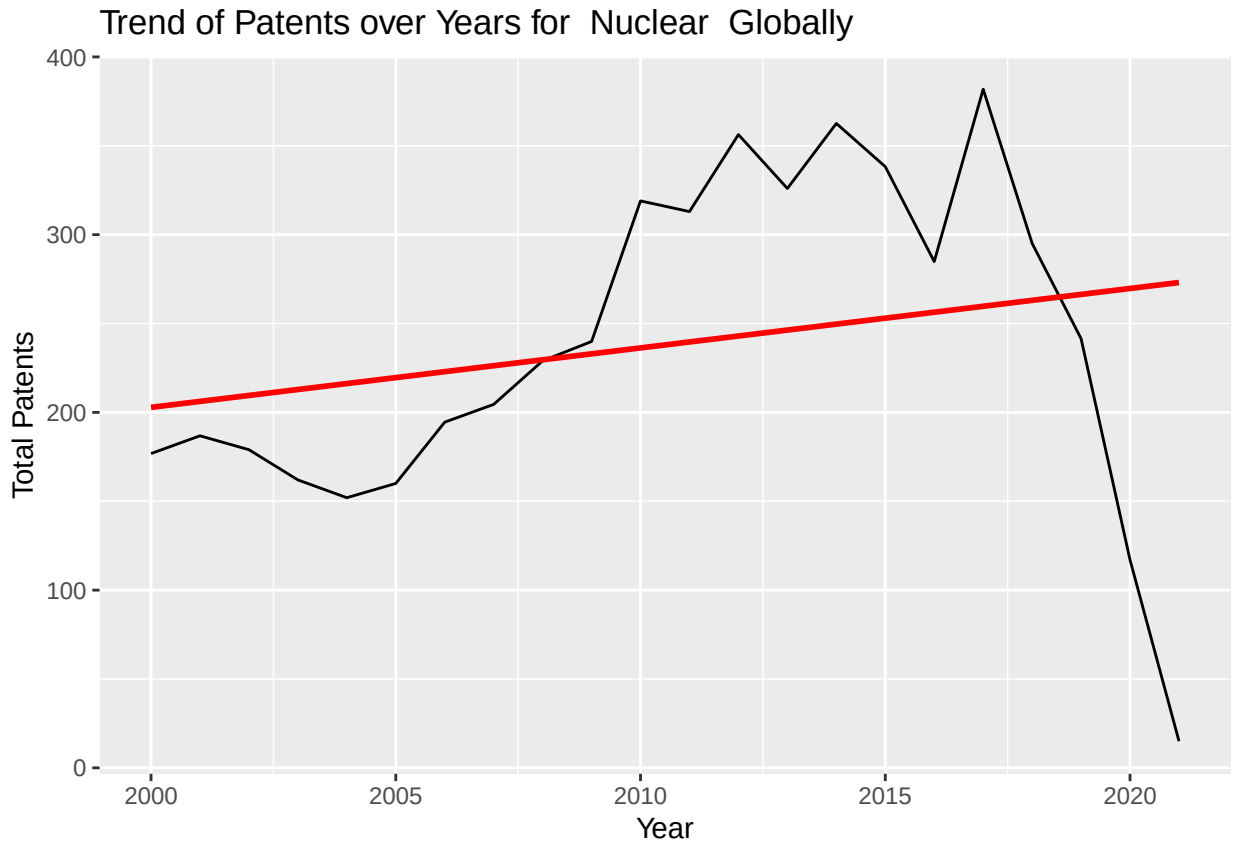

```
# Print the top ten countries
print(clean_energy)
```

```
## # A tibble: 10 x 2
##    countryISO total_value
##    <chr>            <dbl>
##  1 JPN             80478.
##  2 USA             63433.
##  3 CHN             45293.
##  4 DEU             38643.
##  5 KOR             32823.
##  6 FRA             14909.
##  7 GBN              9369.
##  8 GBR              9369.
##  9 TWN              9107.
## 10 SWE              5021.
```

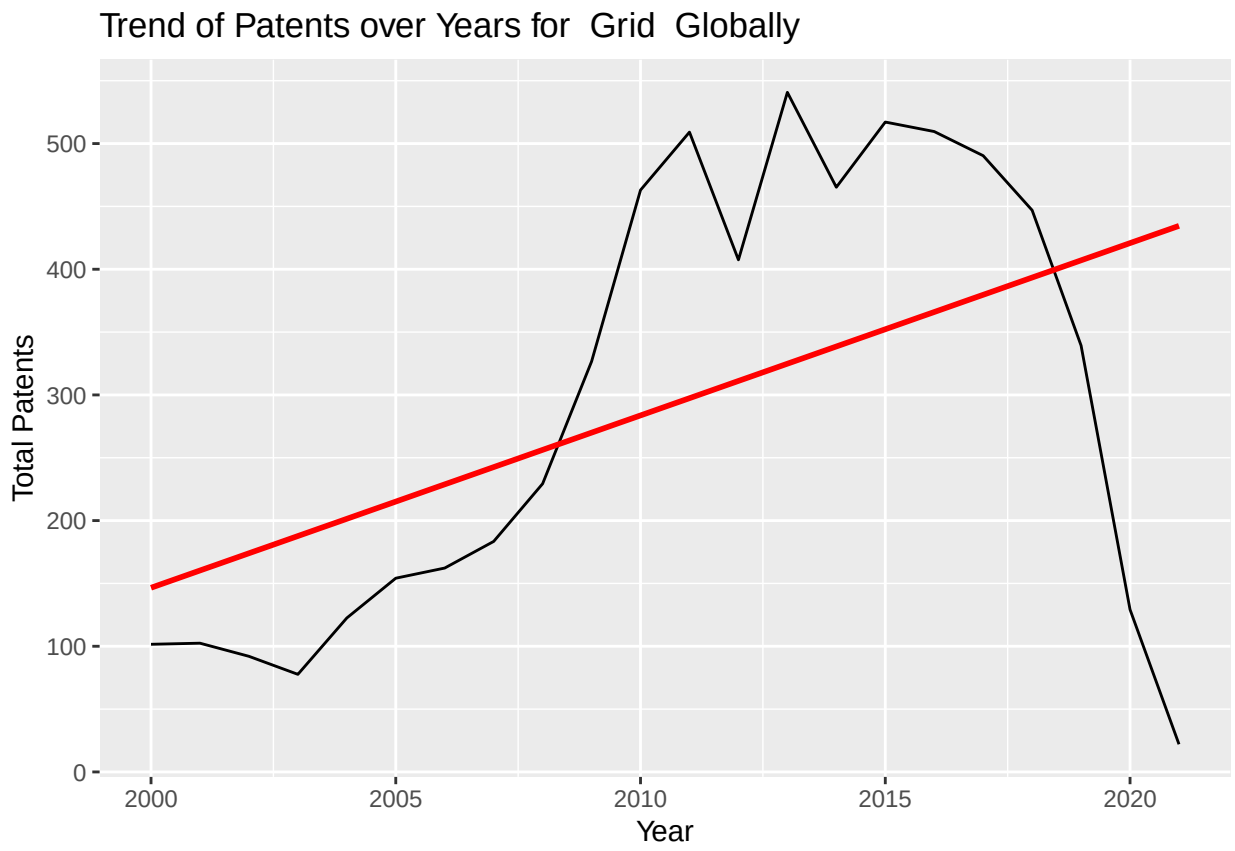
```
ISO <- "WORLD"

for (field in tech_fields) {
  patents_trend_world <- patents_df[grepl(field, patents_df$techSector) & patents_df$countryISO == ISO,

  # Only plot the graph if there is more than one observation for the specific technology field
  if(nrow(patents_trend_world) > 1) {
```

```
    # Summarize the data
    summarized_data <- patents_trend_world %>%
      group_by(year) %>%
      summarise(total_patents = sum(value, na.rm = TRUE))

    # Fit a linear model
    lm_fit <- lm(total_patents ~ year, data = summarized_data)

    # Check if the coefficient for 'year' is positive
    if (coef(lm_fit)[2] > 0) {
      p <- summarized_data %>%
        ggplot(aes(x = year, y = total_patents)) +
        geom_line() +
        geom_smooth(method = "lm", se = FALSE, col = "red") + # Add regression line
        labs(title = paste("Trend of Patents over Years for ", field, " Globally"), x = "Year", y = "To

      print(p)
    }
  }
}
```

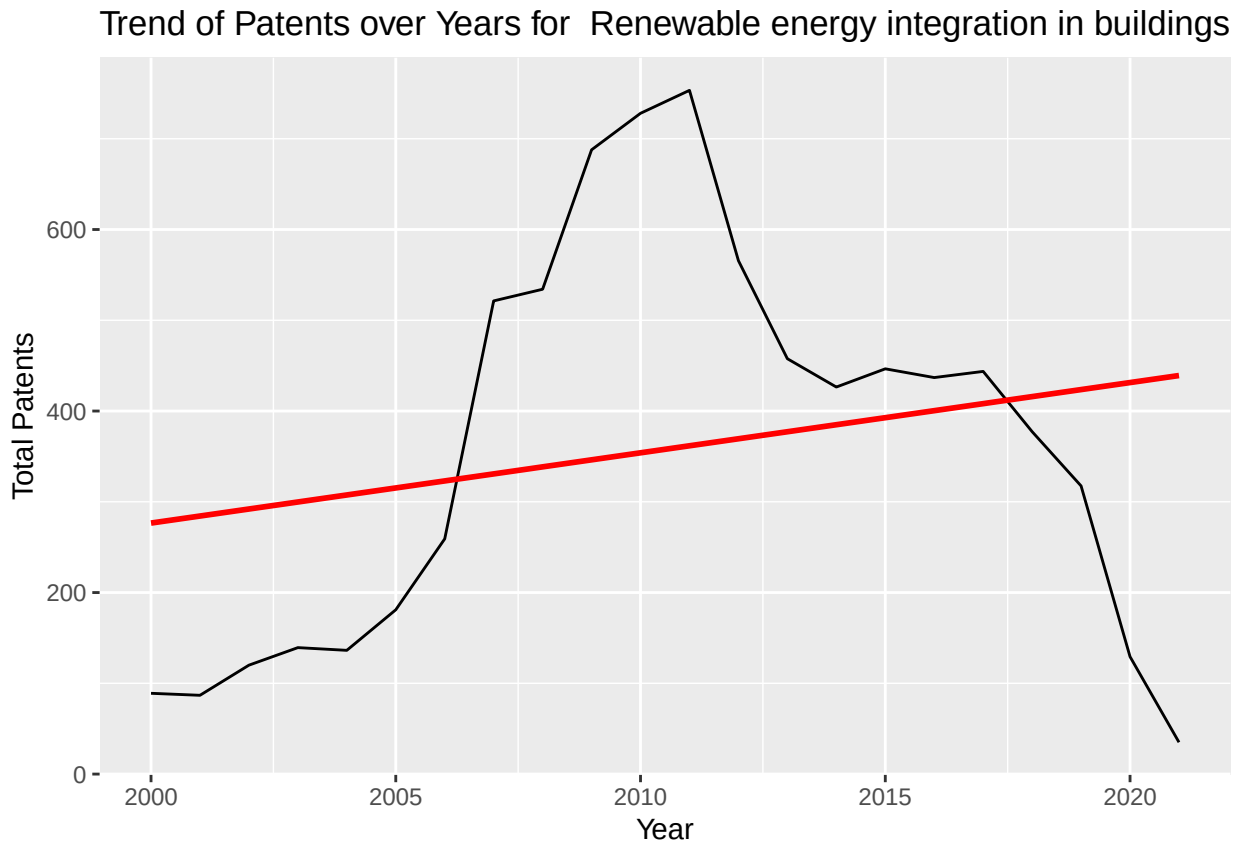
```
## `geom_smooth()` using formula = 'y ~ x'
```



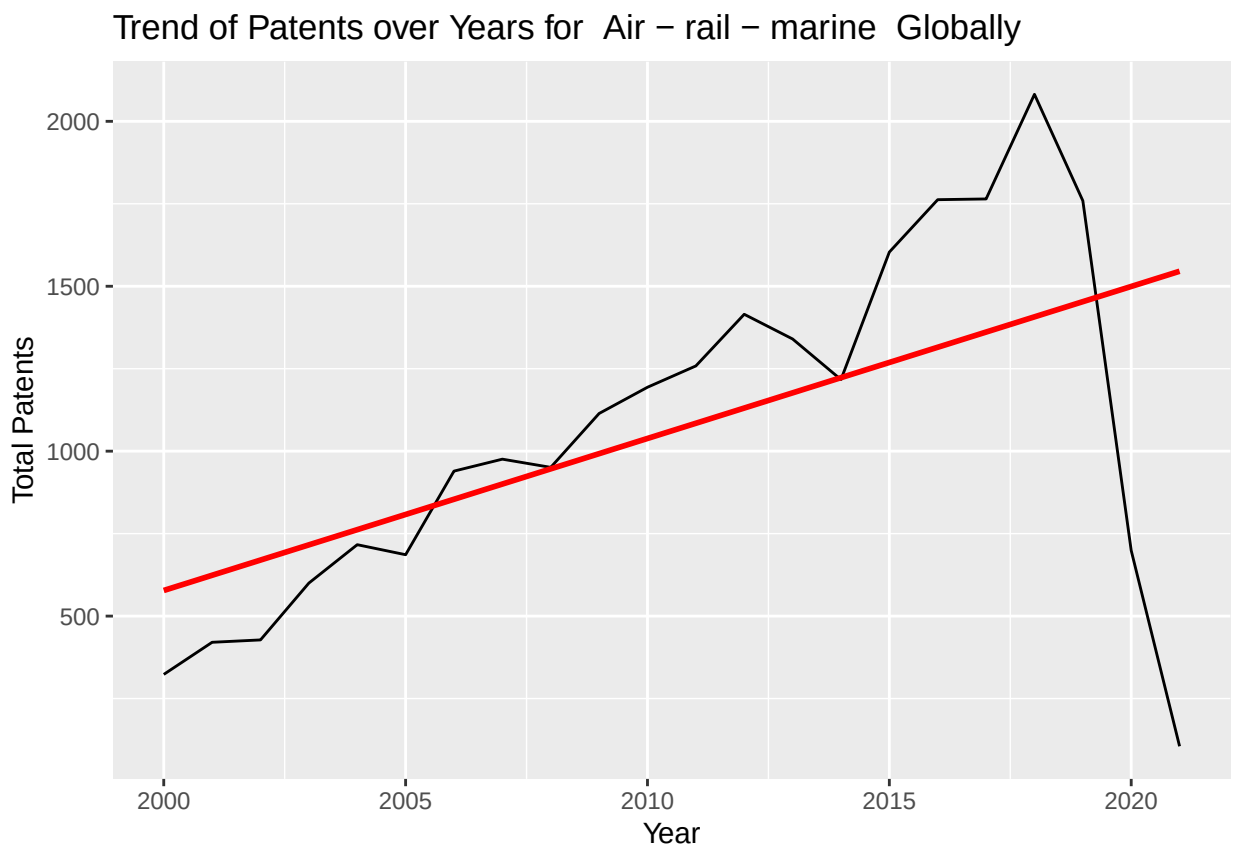

```
## `geom_smooth()` using formula = 'y ~ x'
```



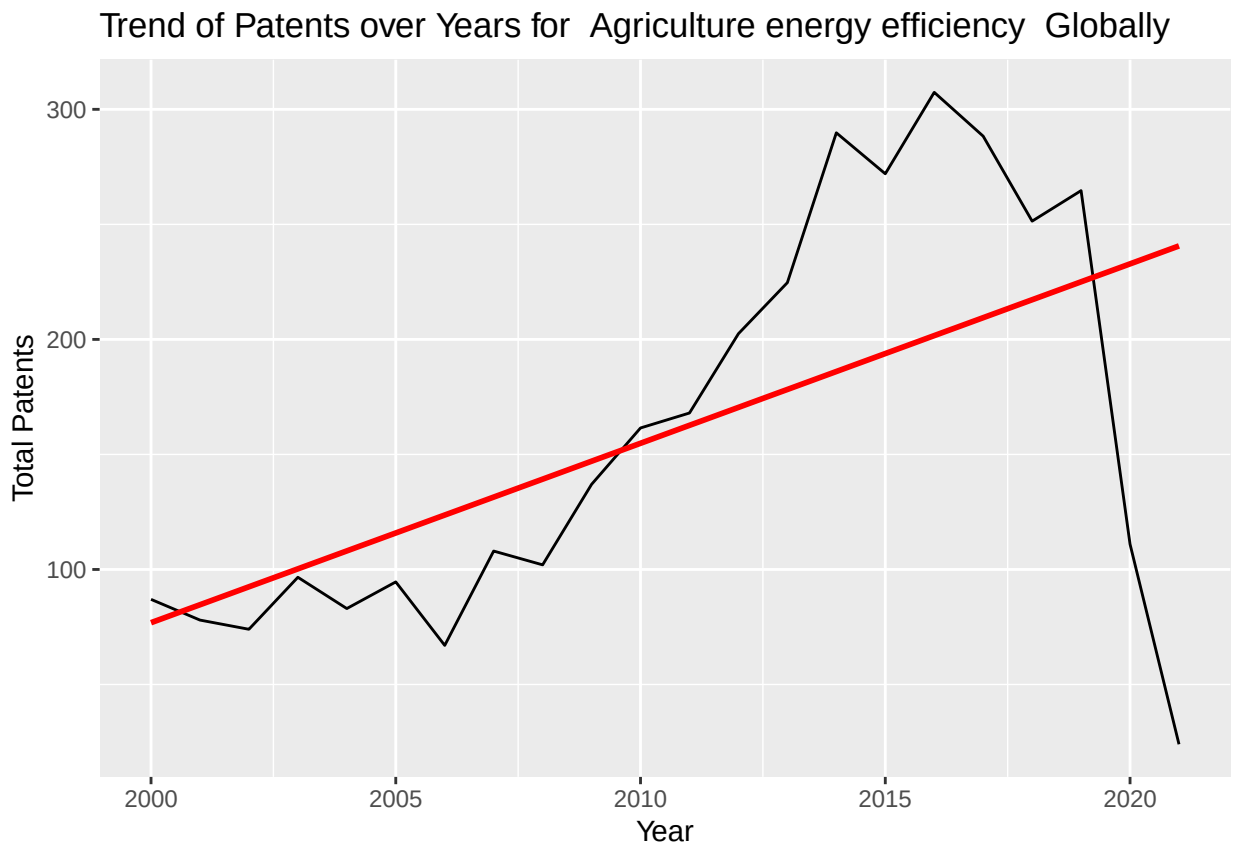

```
## `geom_smooth()` using formula = 'y ~ x'
```



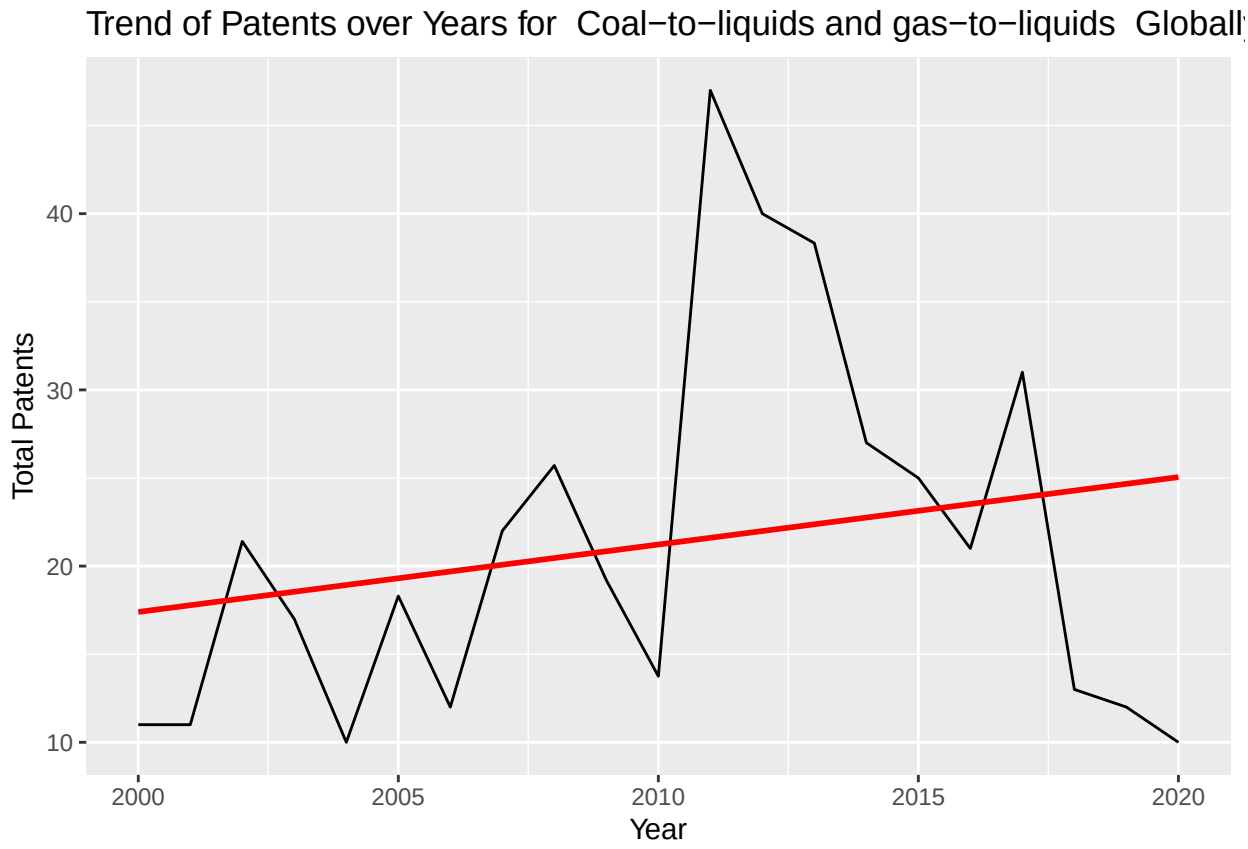

```
## `geom_smooth()` using formula = 'y ~ x'
```



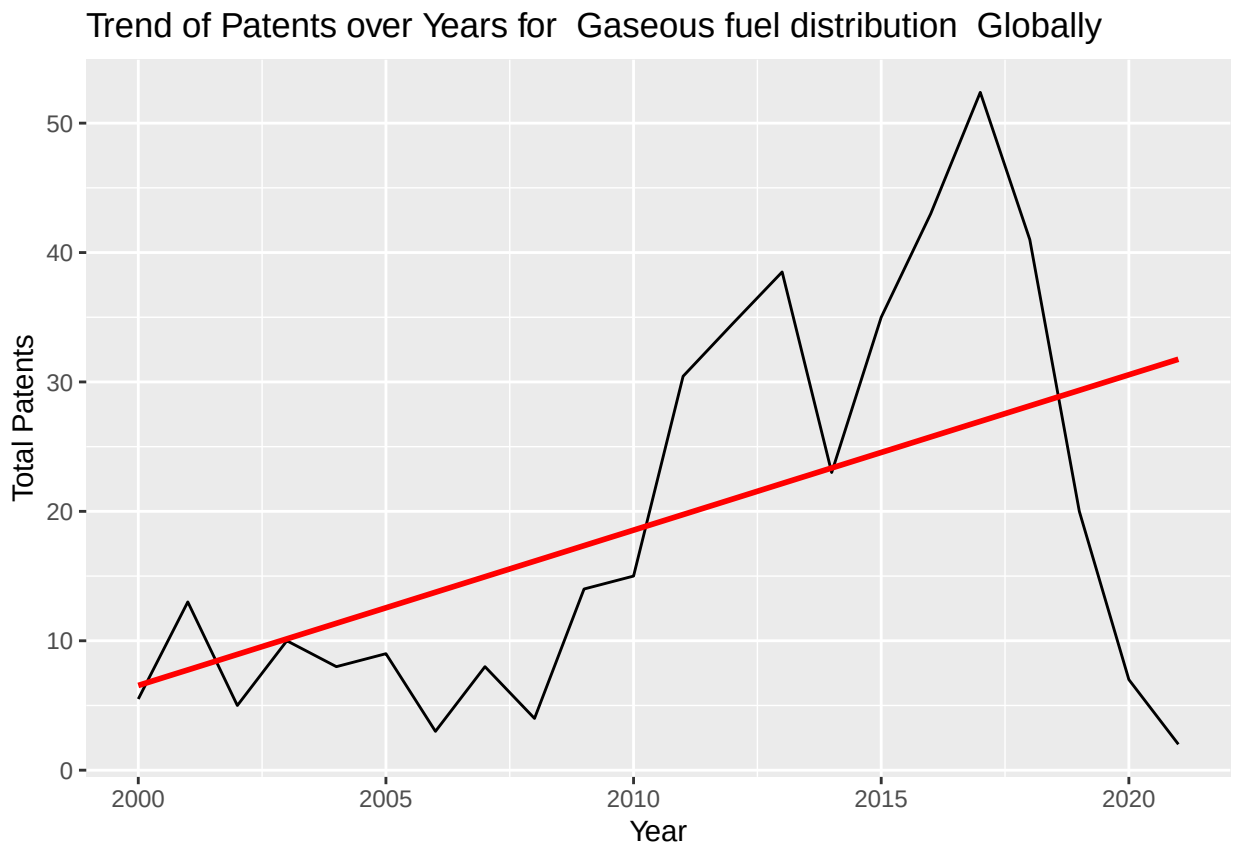

```
## `geom_smooth()` using formula = 'y ~ x'
```



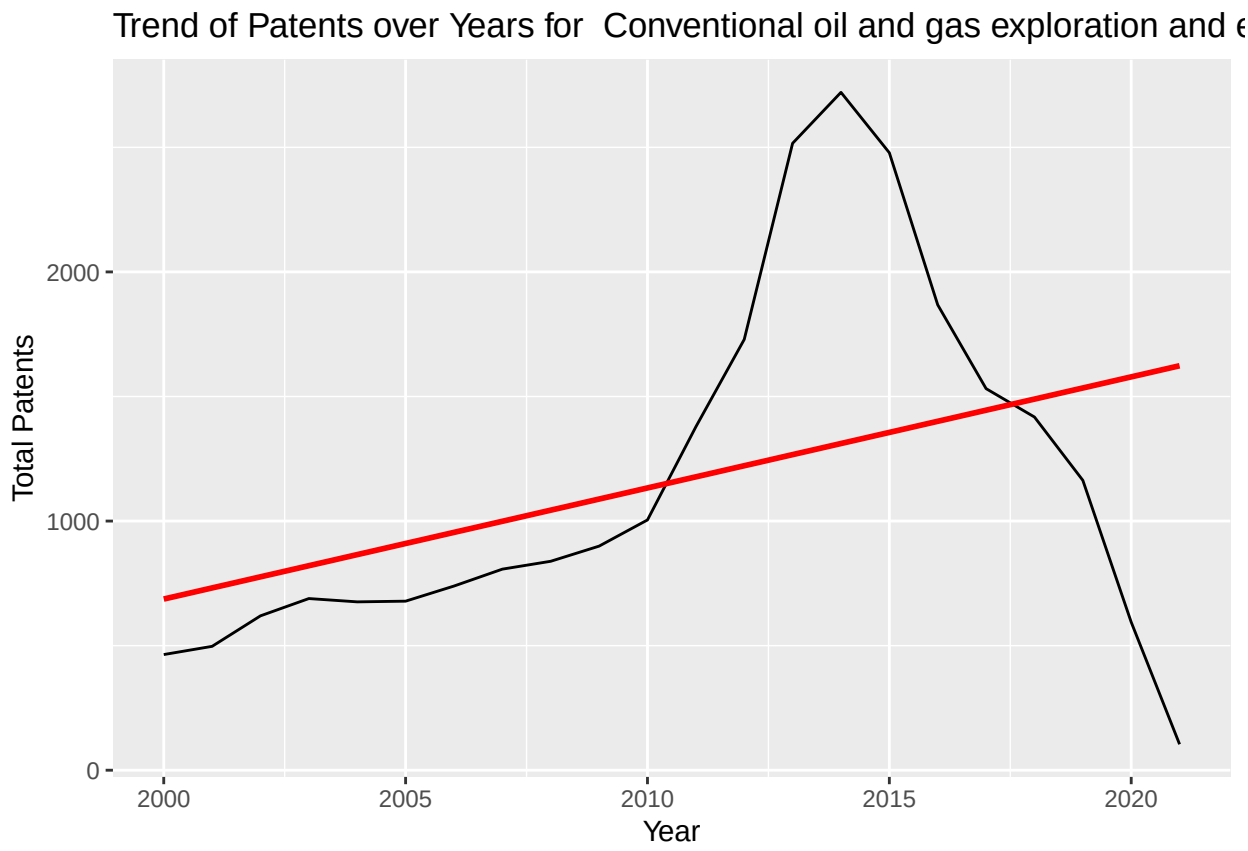

```
## `geom_smooth()` using formula = 'y ~ x'
```



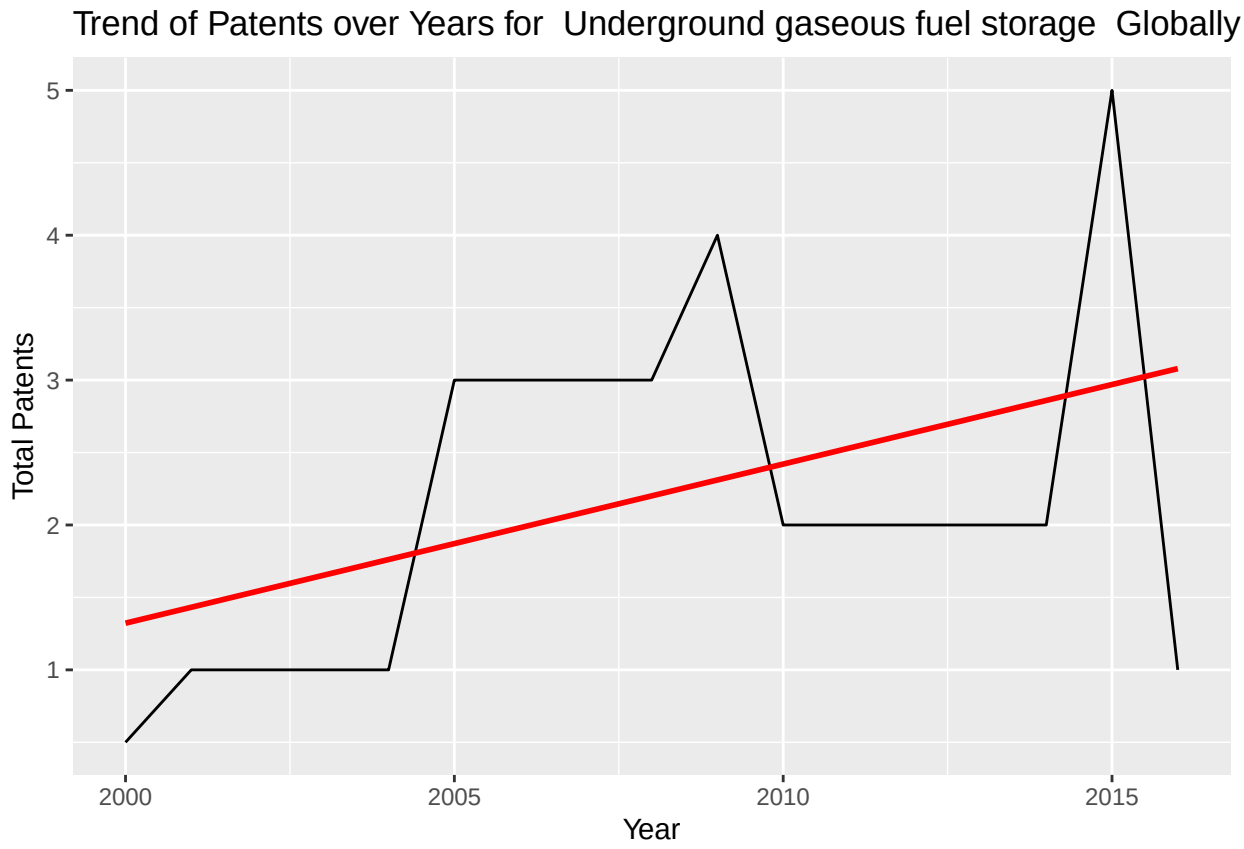

```
## `geom_smooth()` using formula = 'y ~ x'
```



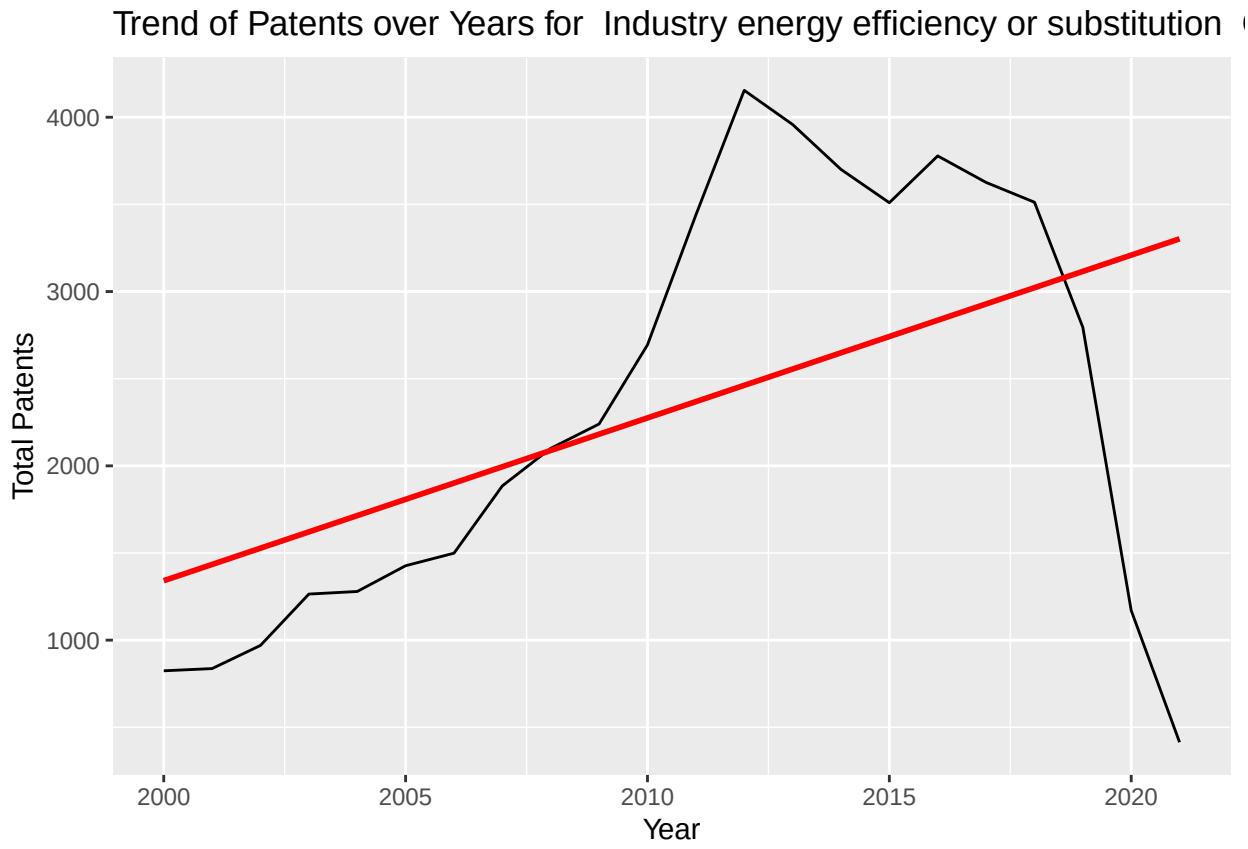

```
## `geom_smooth()` using formula = 'y ~ x'
```



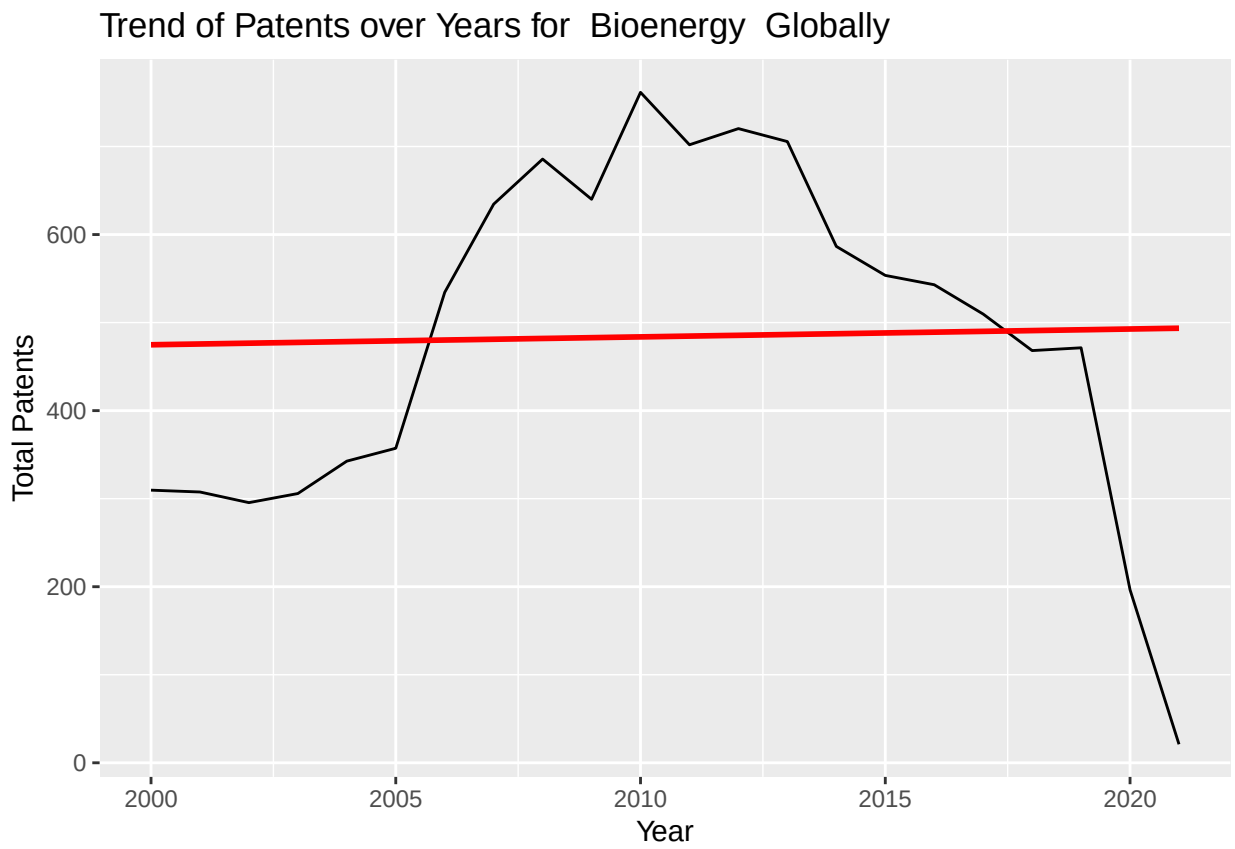

```
## `geom_smooth()` using formula = 'y ~ x'
```



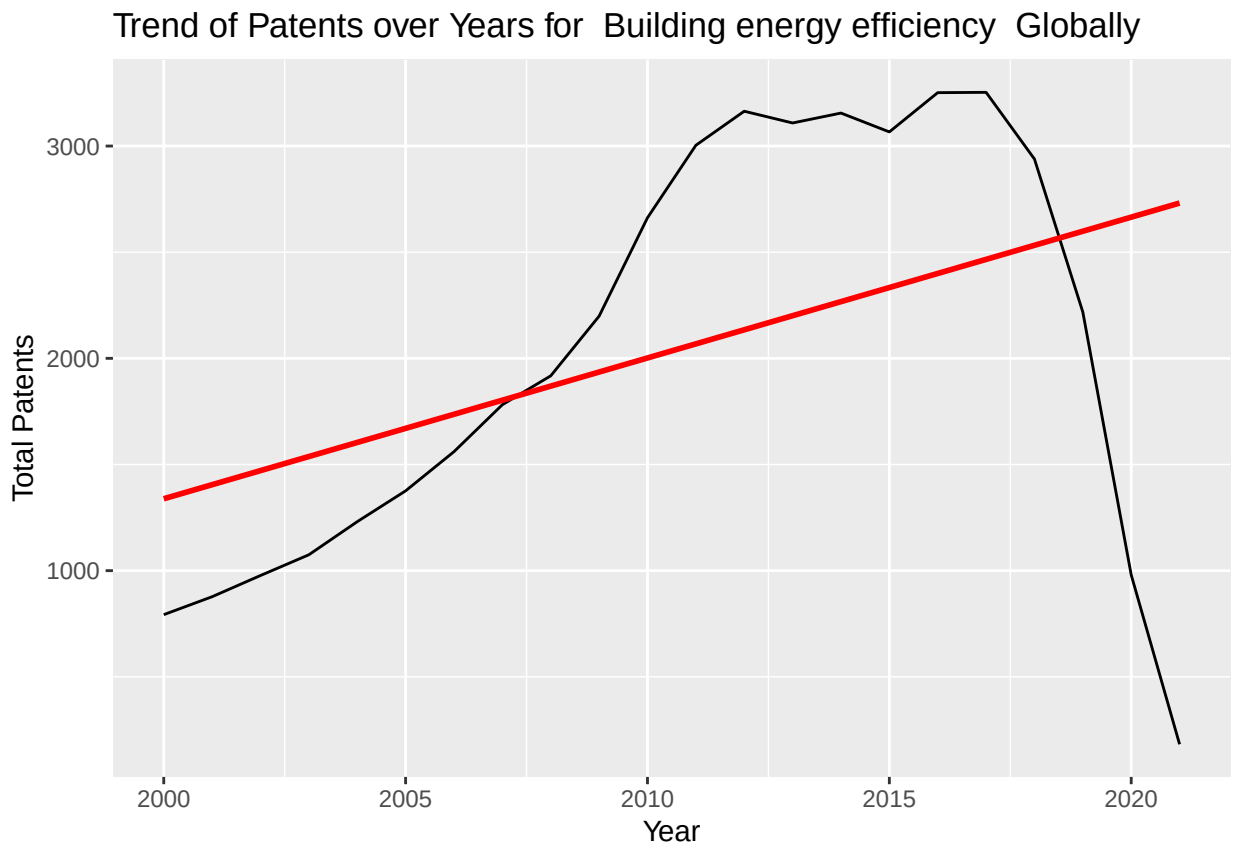

```
## `geom_smooth()` using formula = 'y ~ x'
```



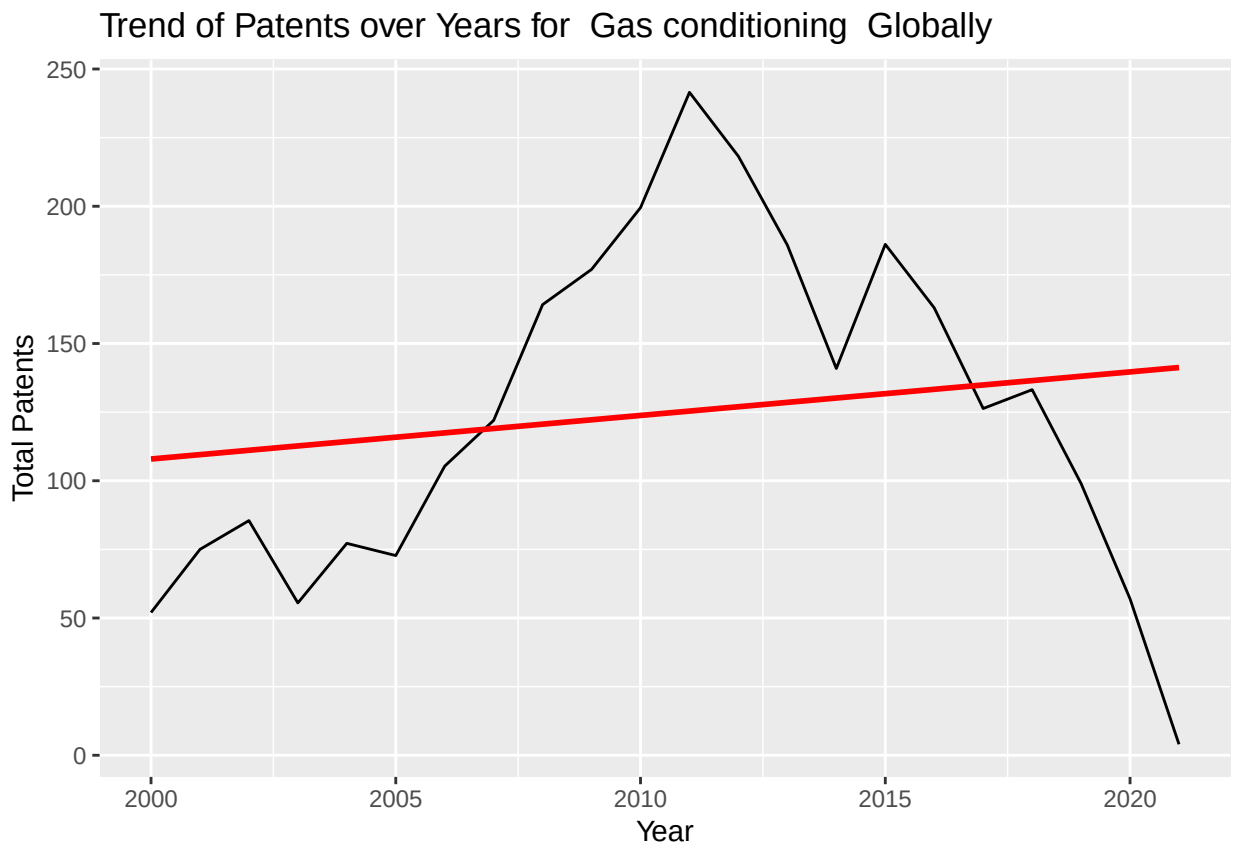

```
## `geom_smooth()` using formula = 'y ~ x'
```



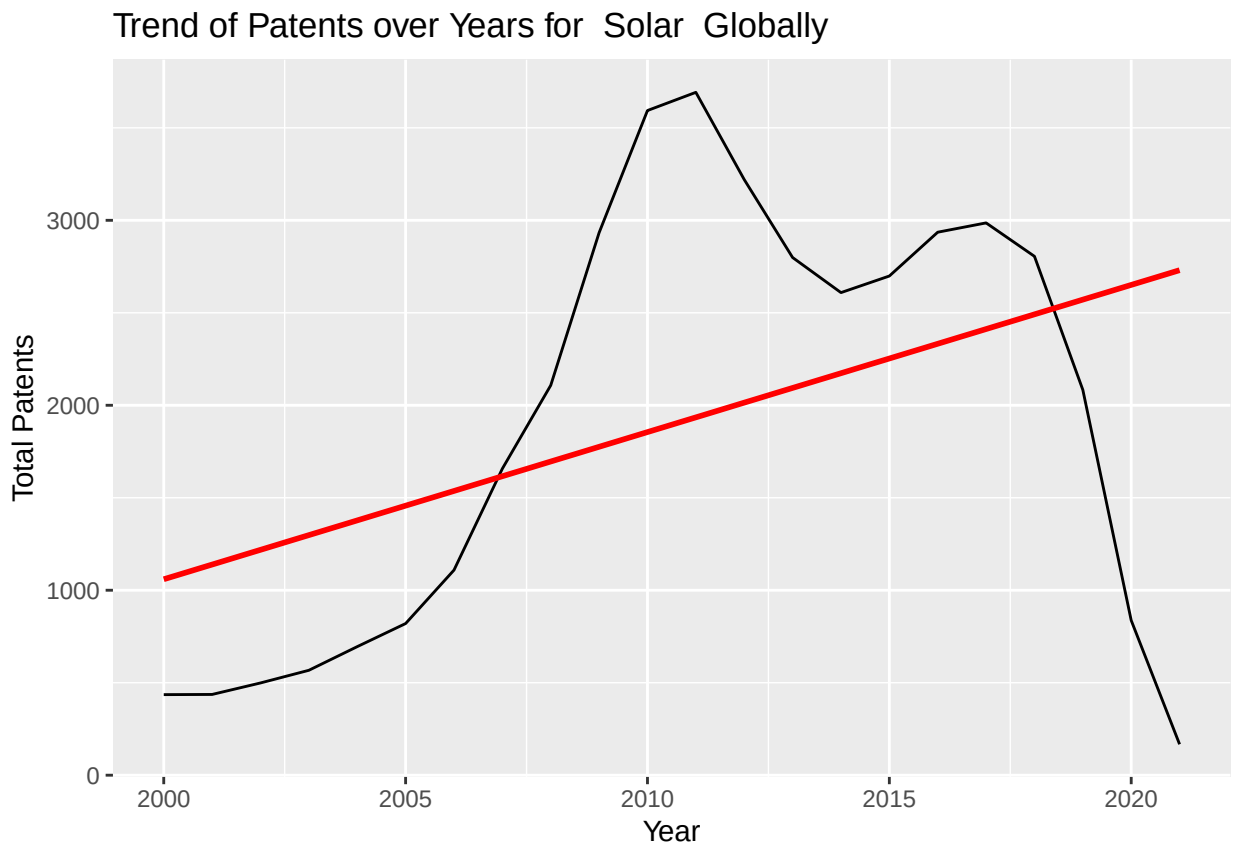

```
## `geom_smooth()` using formula = 'y ~ x'
```



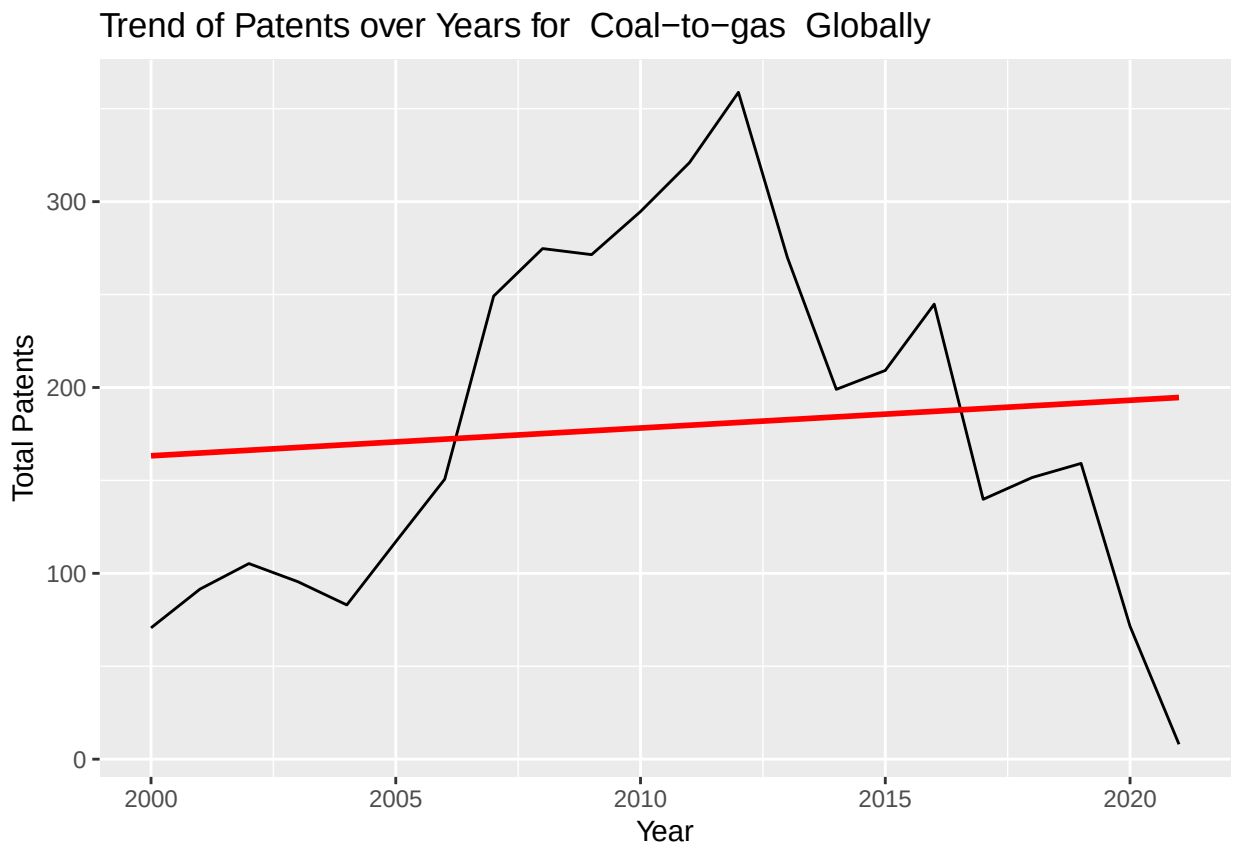

```
## `geom_smooth()` using formula = 'y ~ x'
```



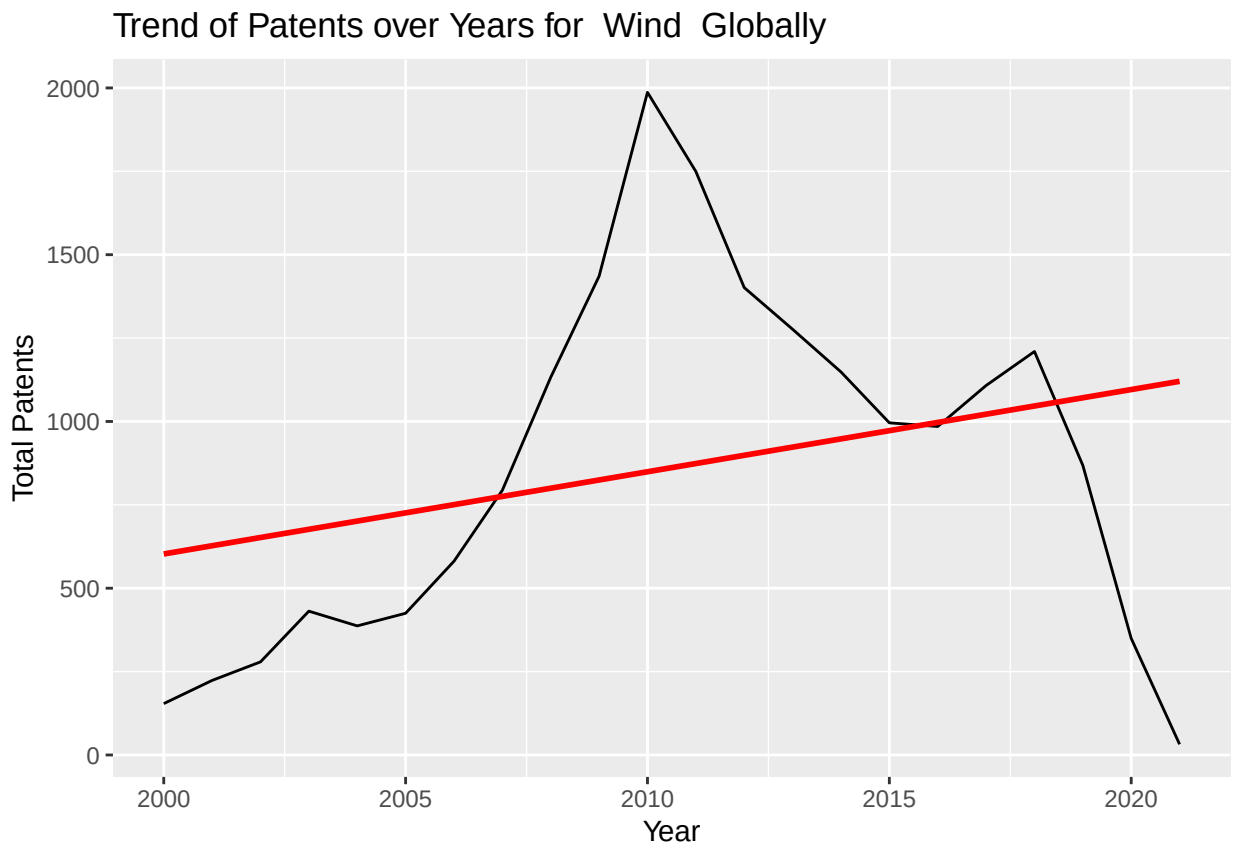

```
## `geom_smooth()` using formula = 'y ~ x'
```



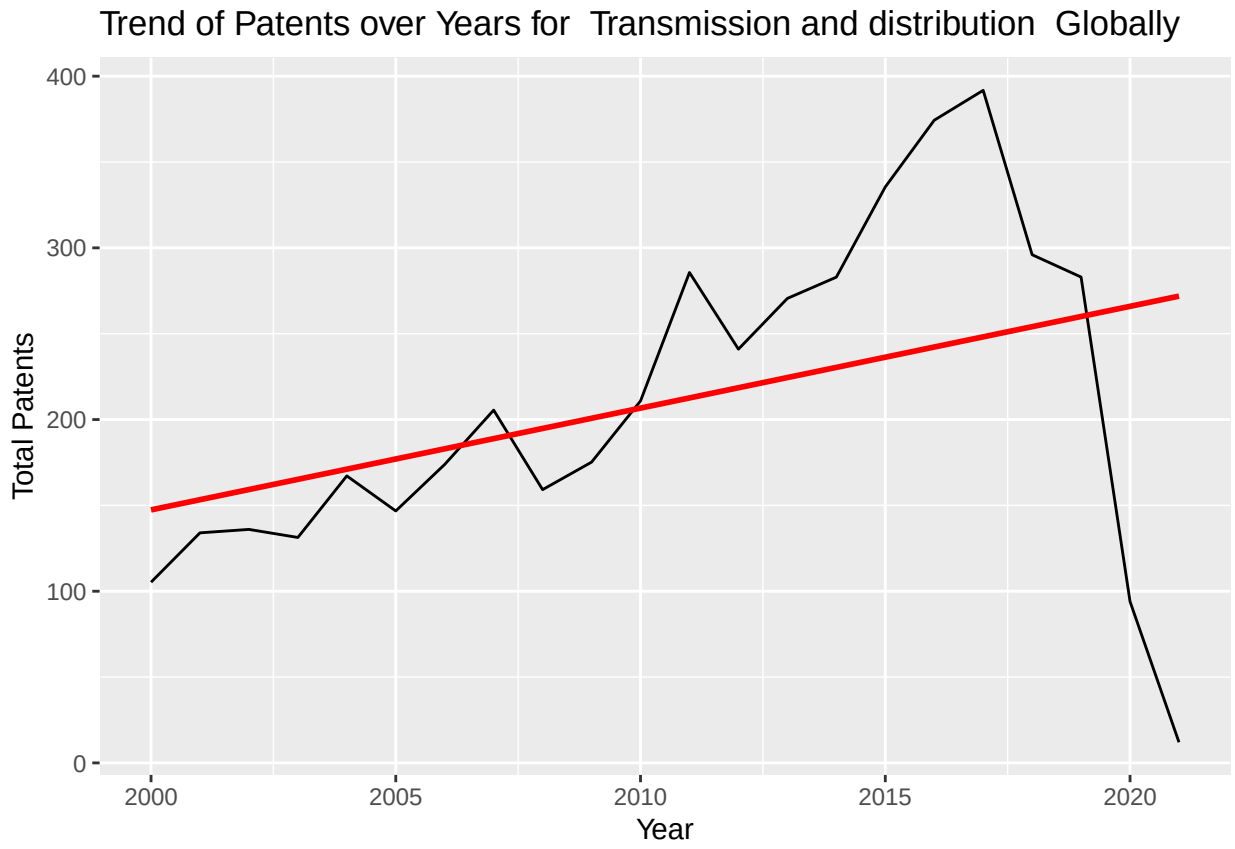

```
## `geom_smooth()` using formula = 'y ~ x'
```



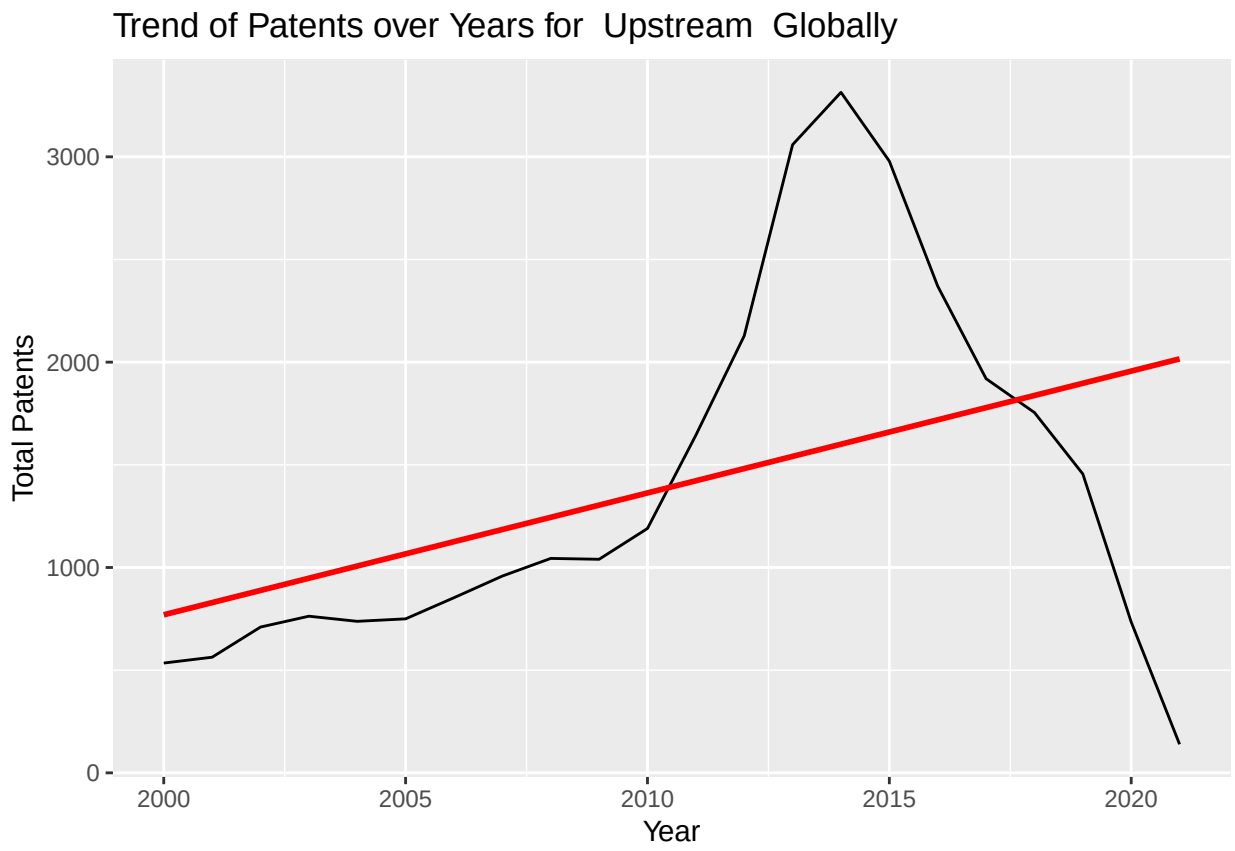

```
## `geom_smooth()` using formula = 'y ~ x'
```



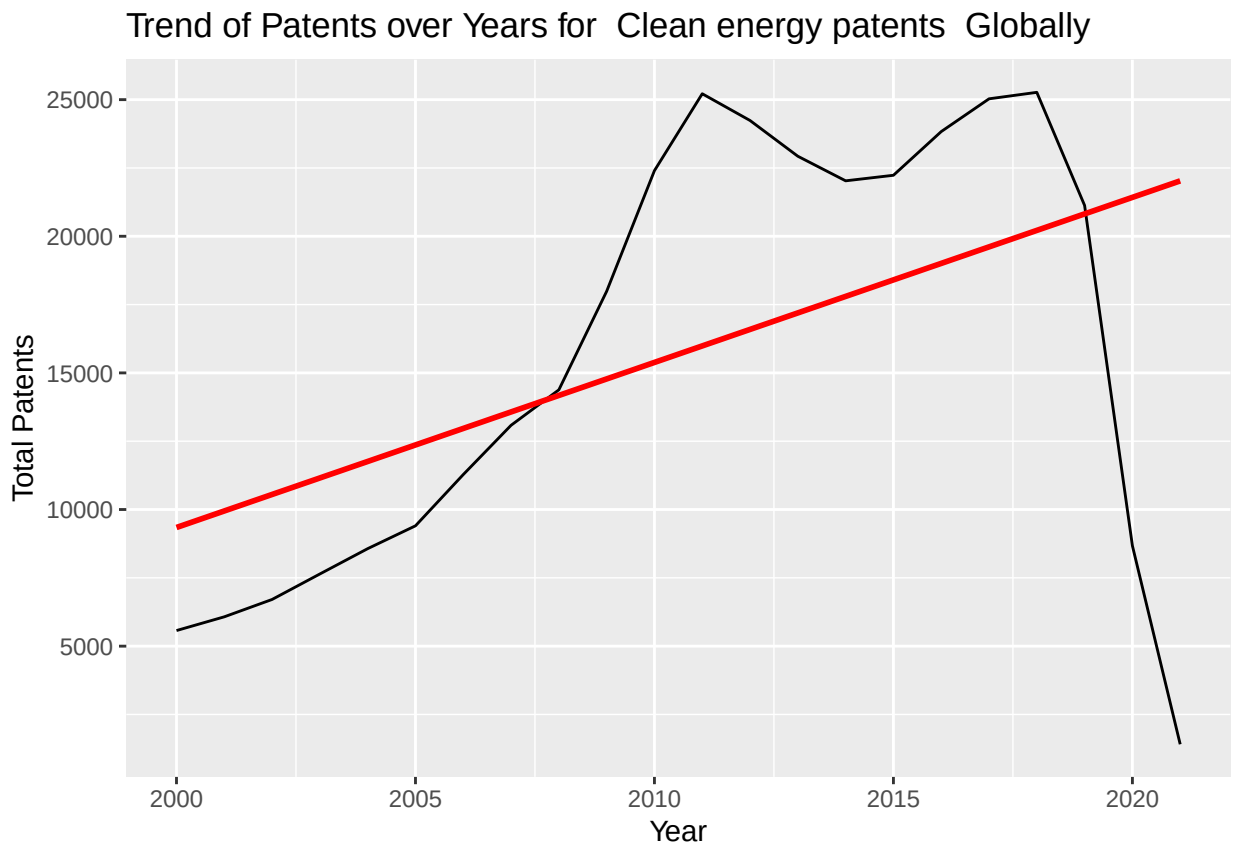

```
## `geom_smooth()` using formula = 'y ~ x'
```



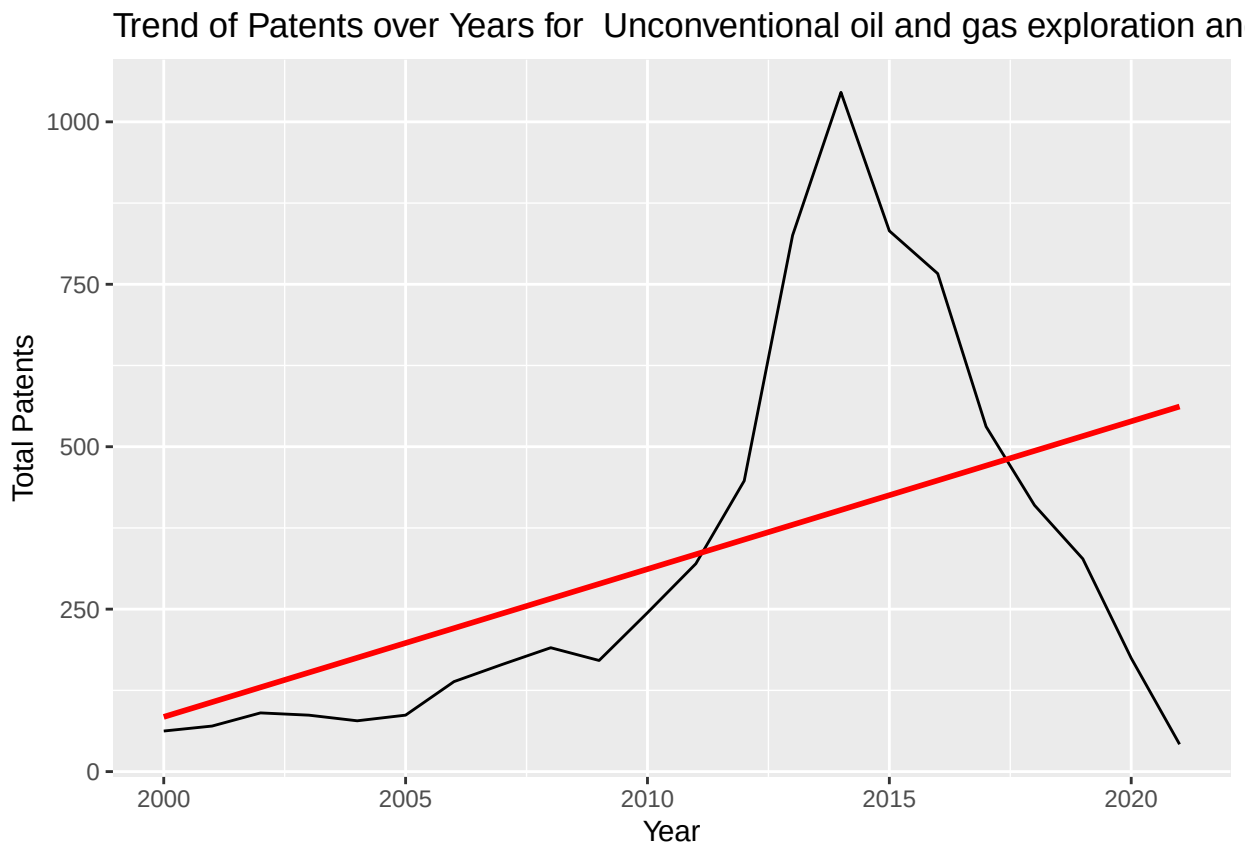

```
## `geom_smooth()` using formula = 'y ~ x'
```



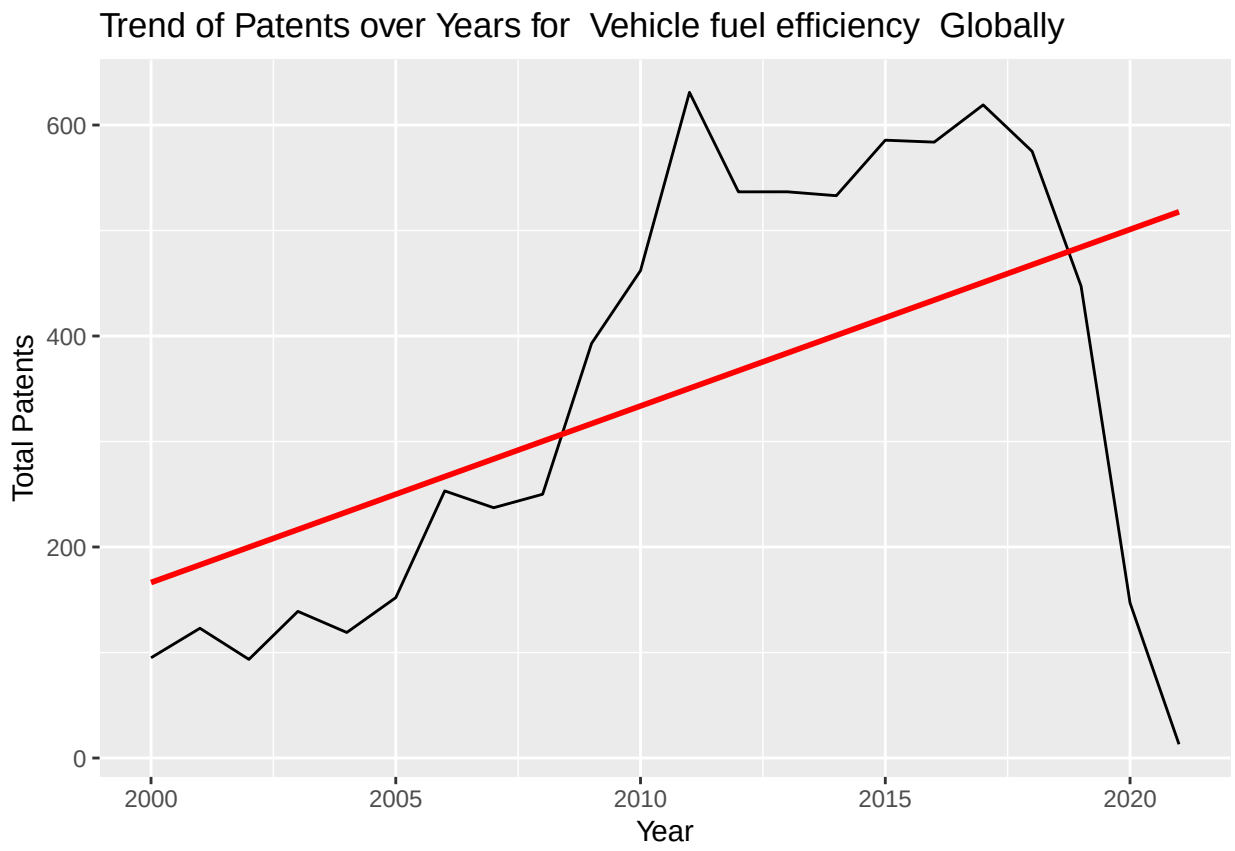

```
## `geom_smooth()` using formula = 'y ~ x'
```



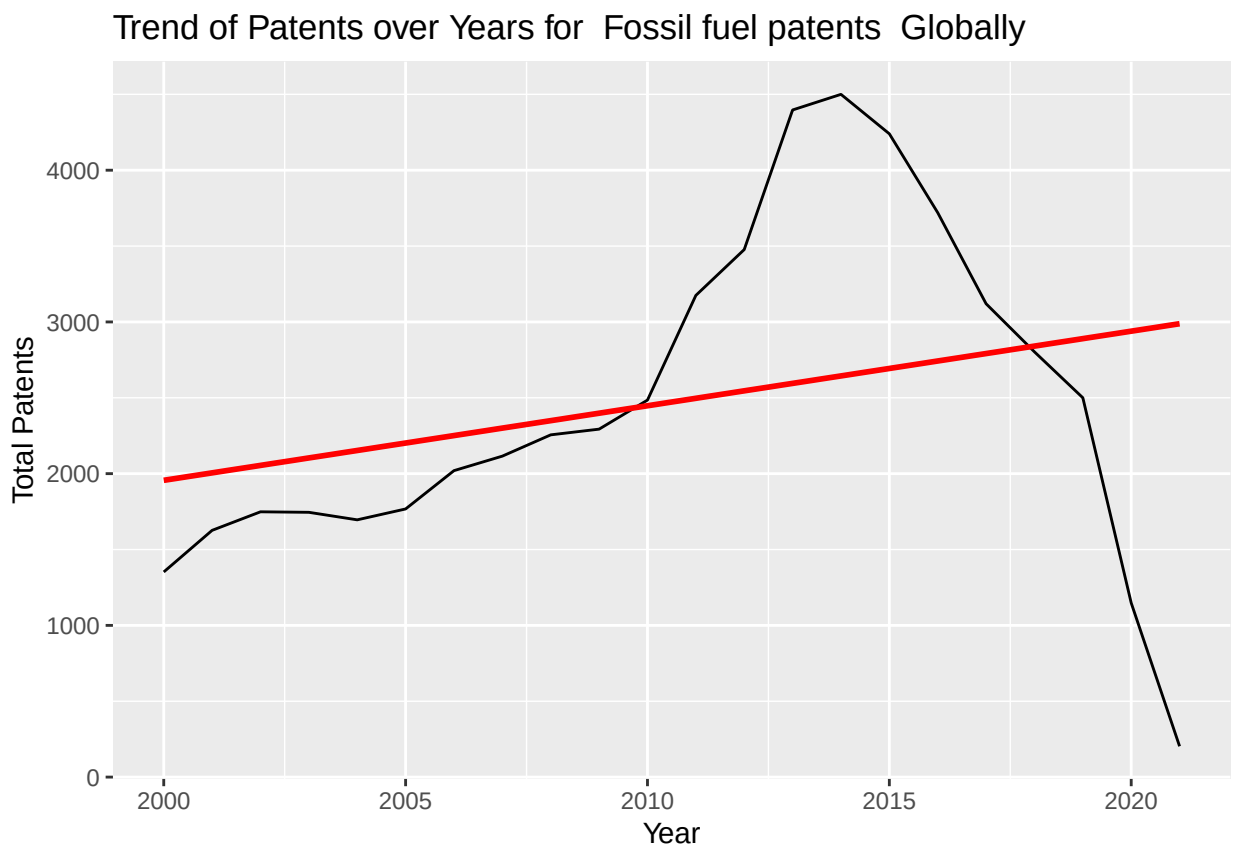

```
## `geom_smooth()` using formula = 'y ~ x'
```



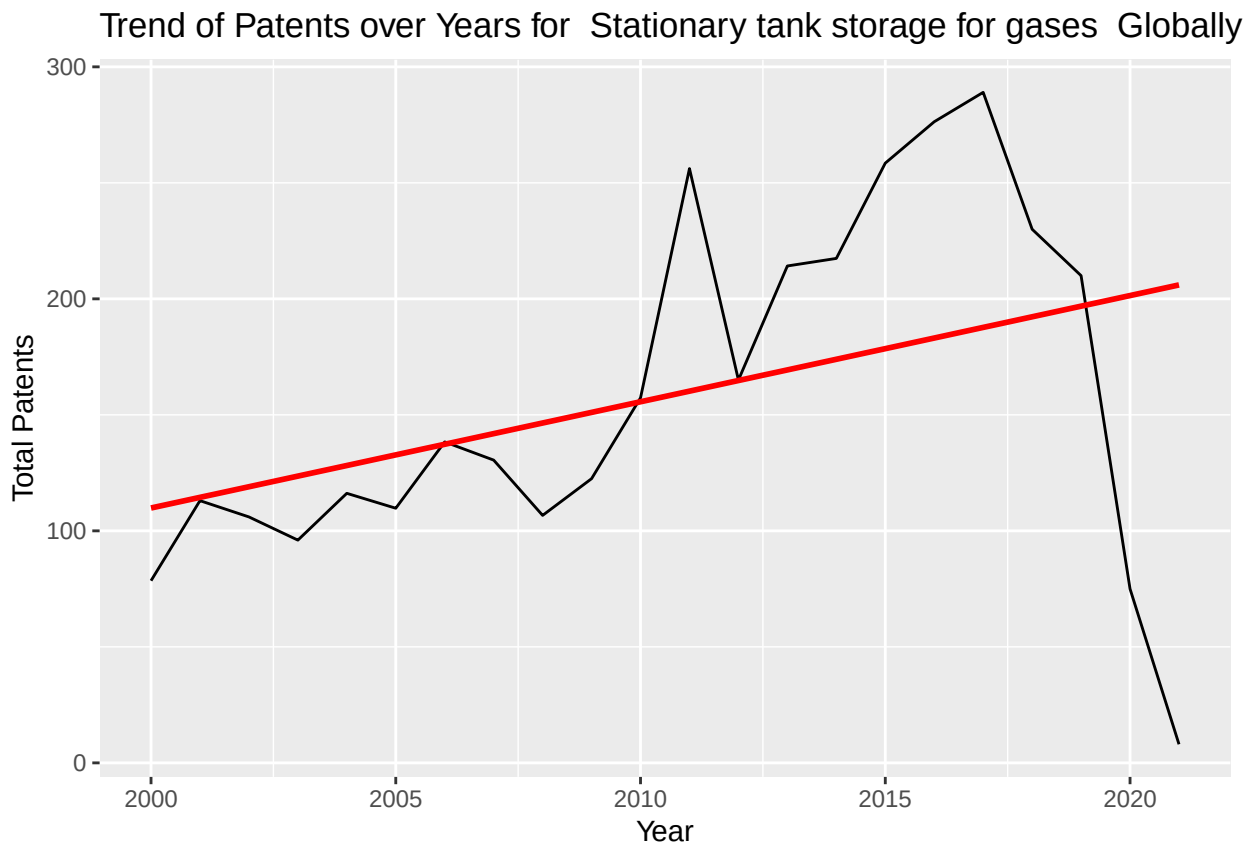

```
## `geom_smooth()` using formula = 'y ~ x'
```



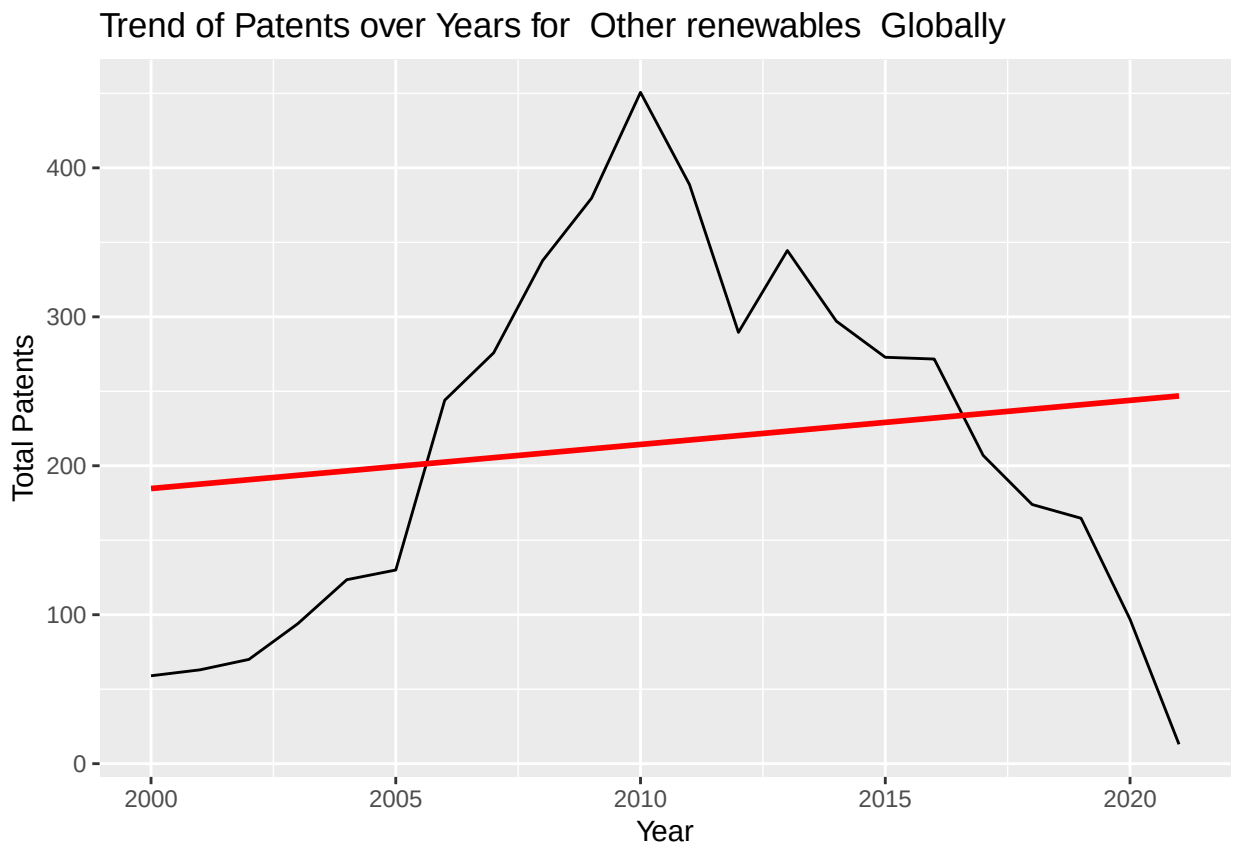

```
## `geom_smooth()` using formula = 'y ~ x'
```



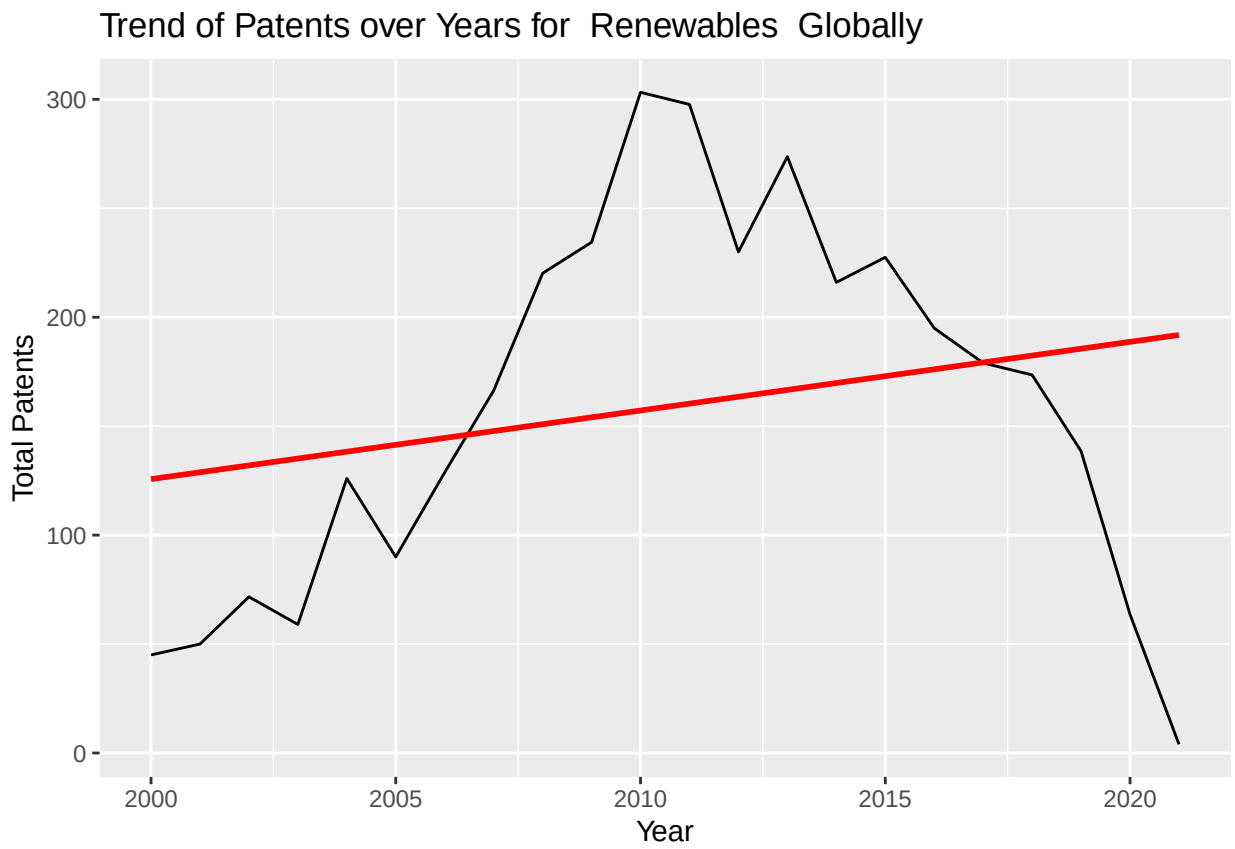

```
## `geom_smooth()` using formula = 'y ~ x'
```



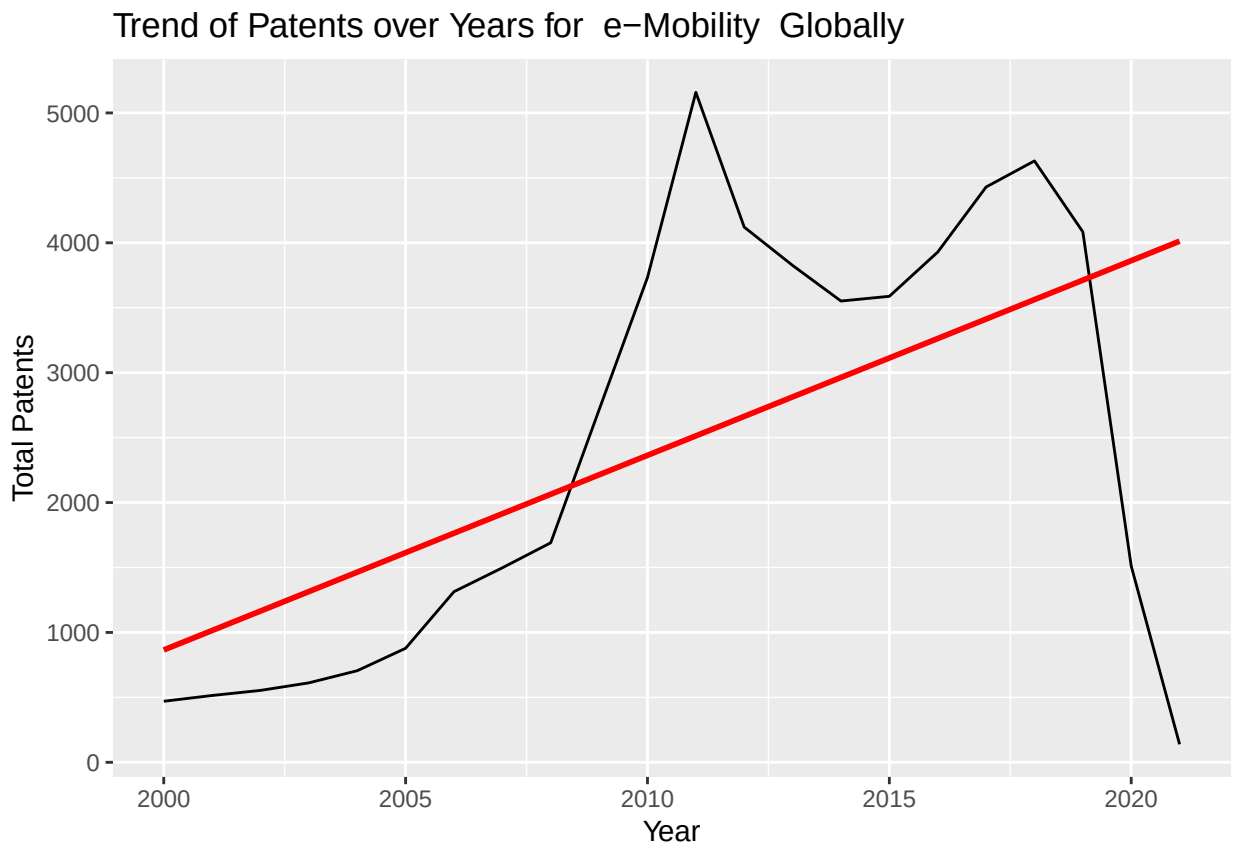

```
## `geom_smooth()` using formula = 'y ~ x'
```



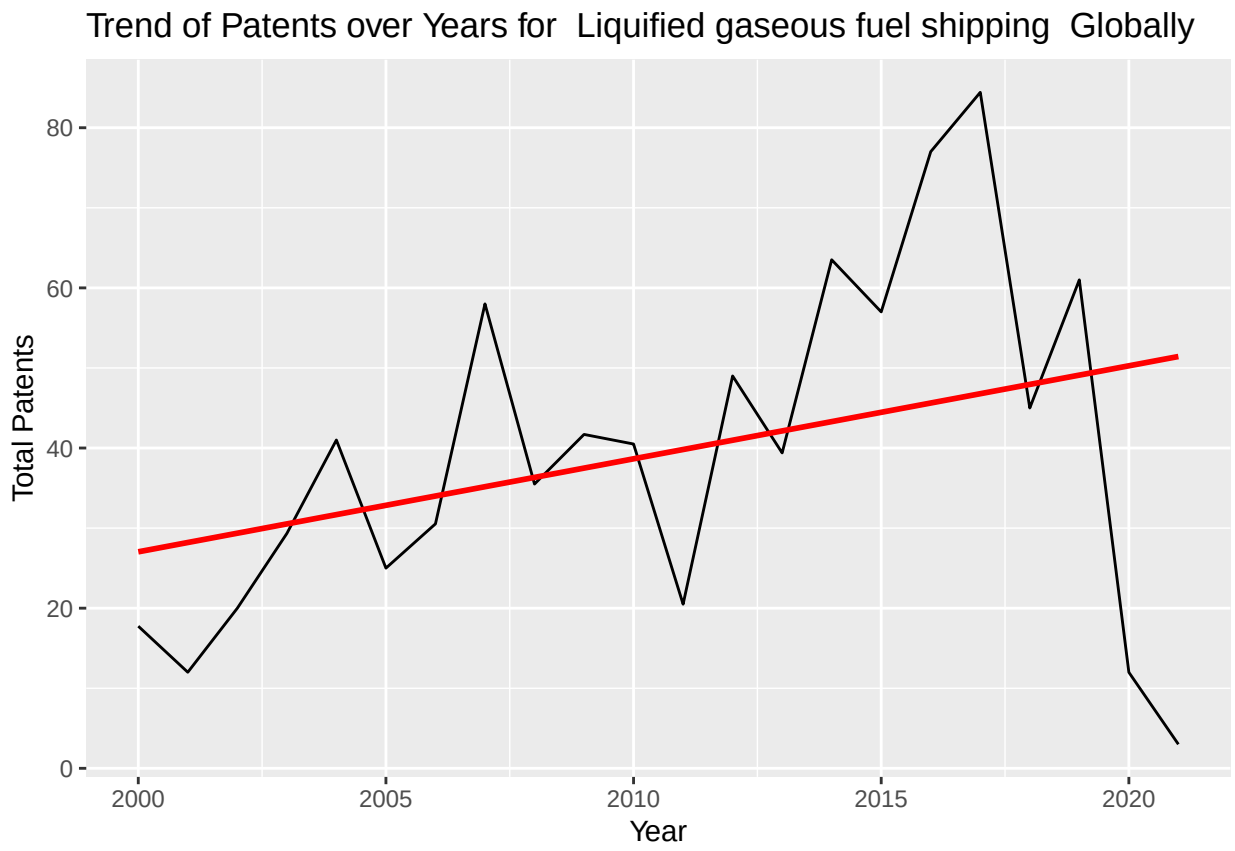

```
## `geom_smooth()` using formula = 'y ~ x'
```



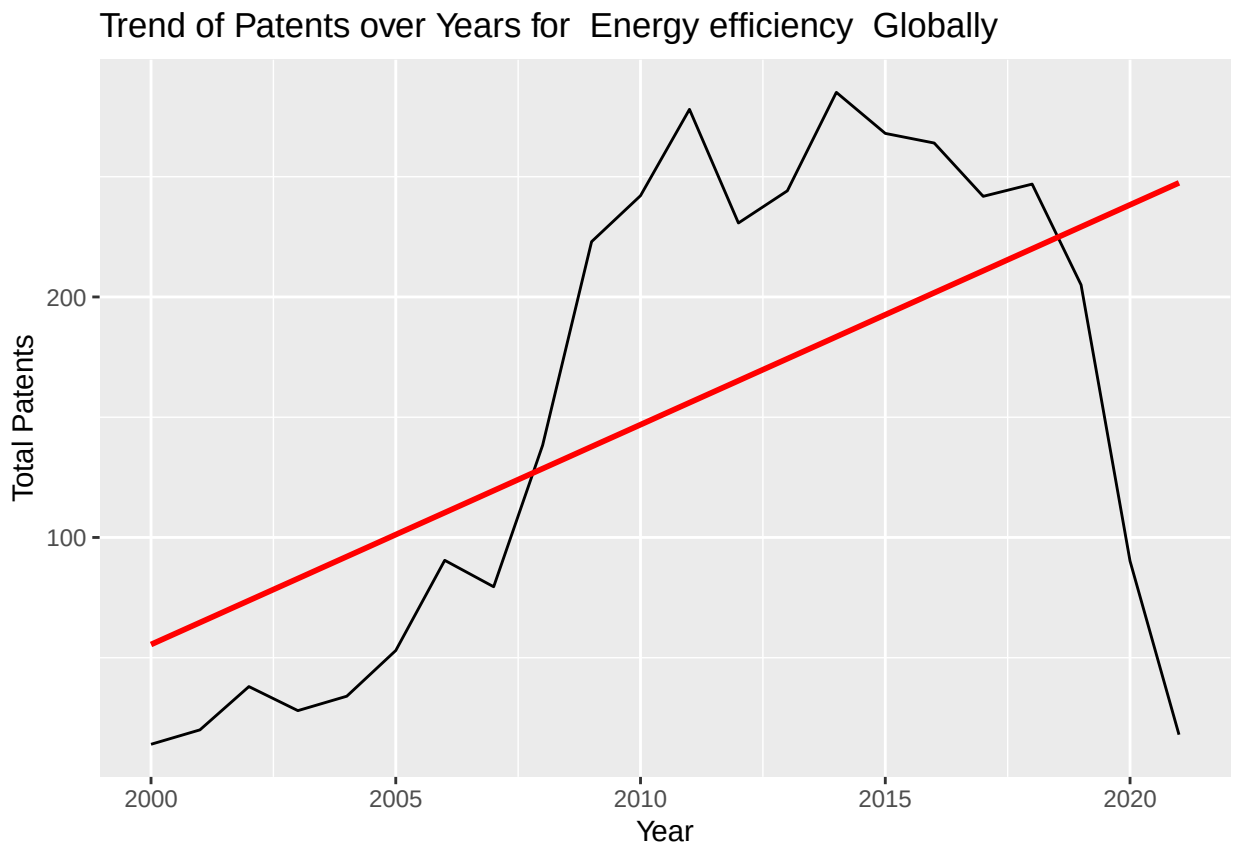

```
## `geom_smooth()` using formula = 'y ~ x'
```



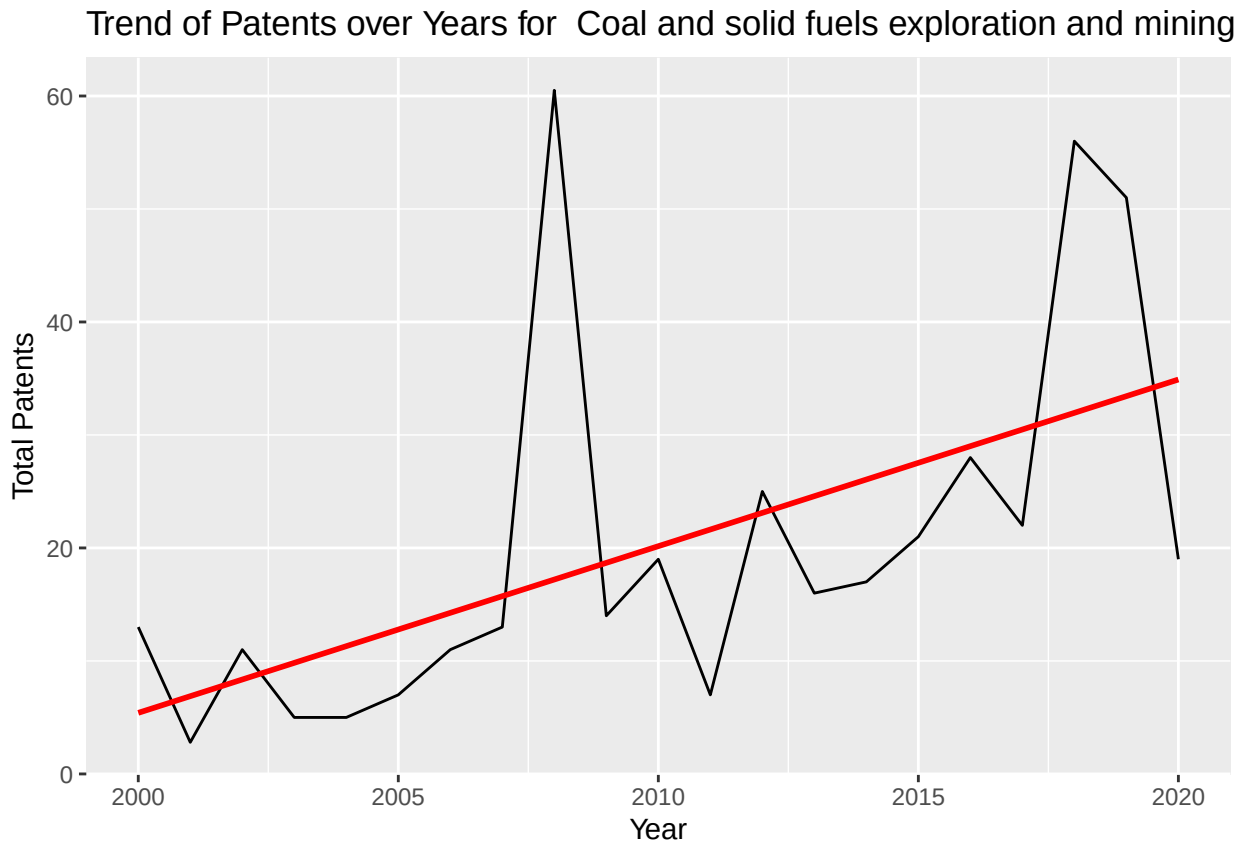

```
## `geom_smooth()` using formula = 'y ~ x'
```



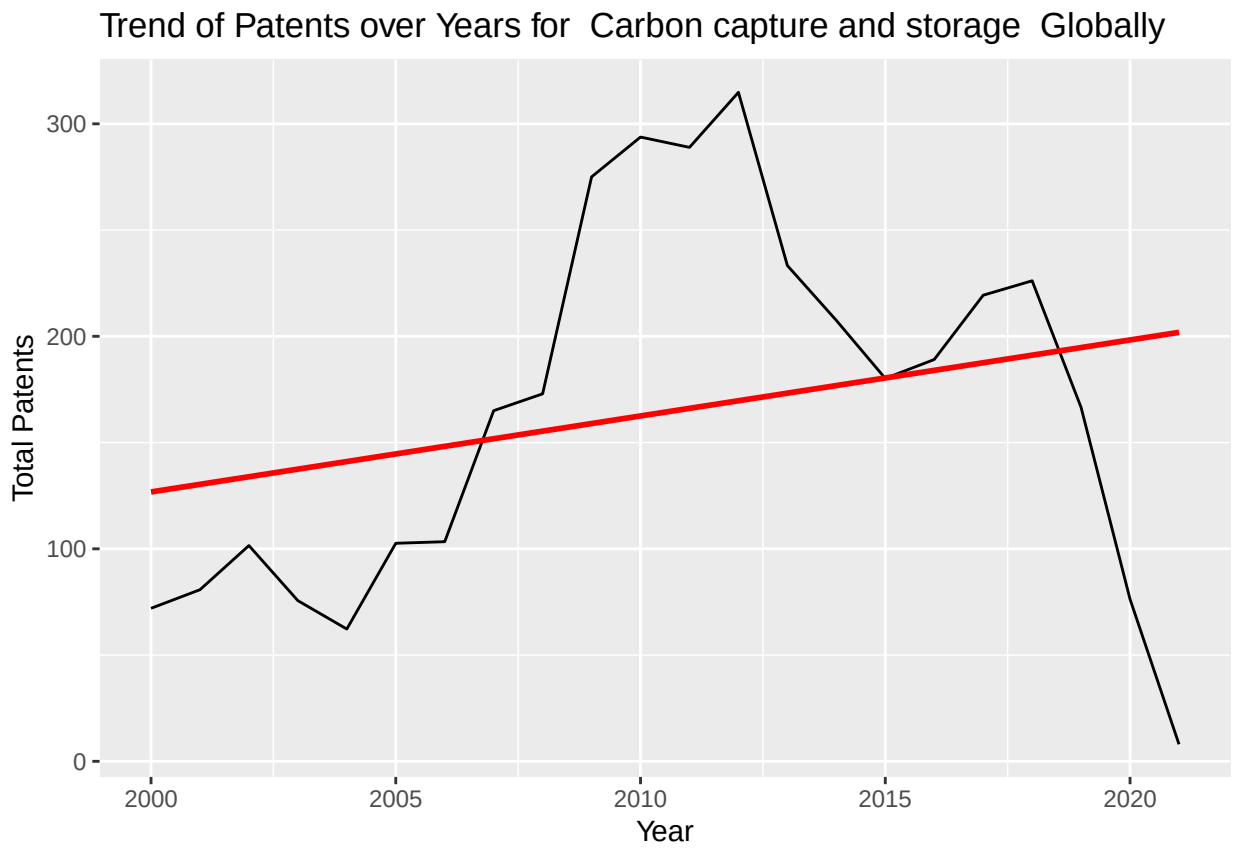

```
## `geom_smooth()` using formula = 'y ~ x'
```



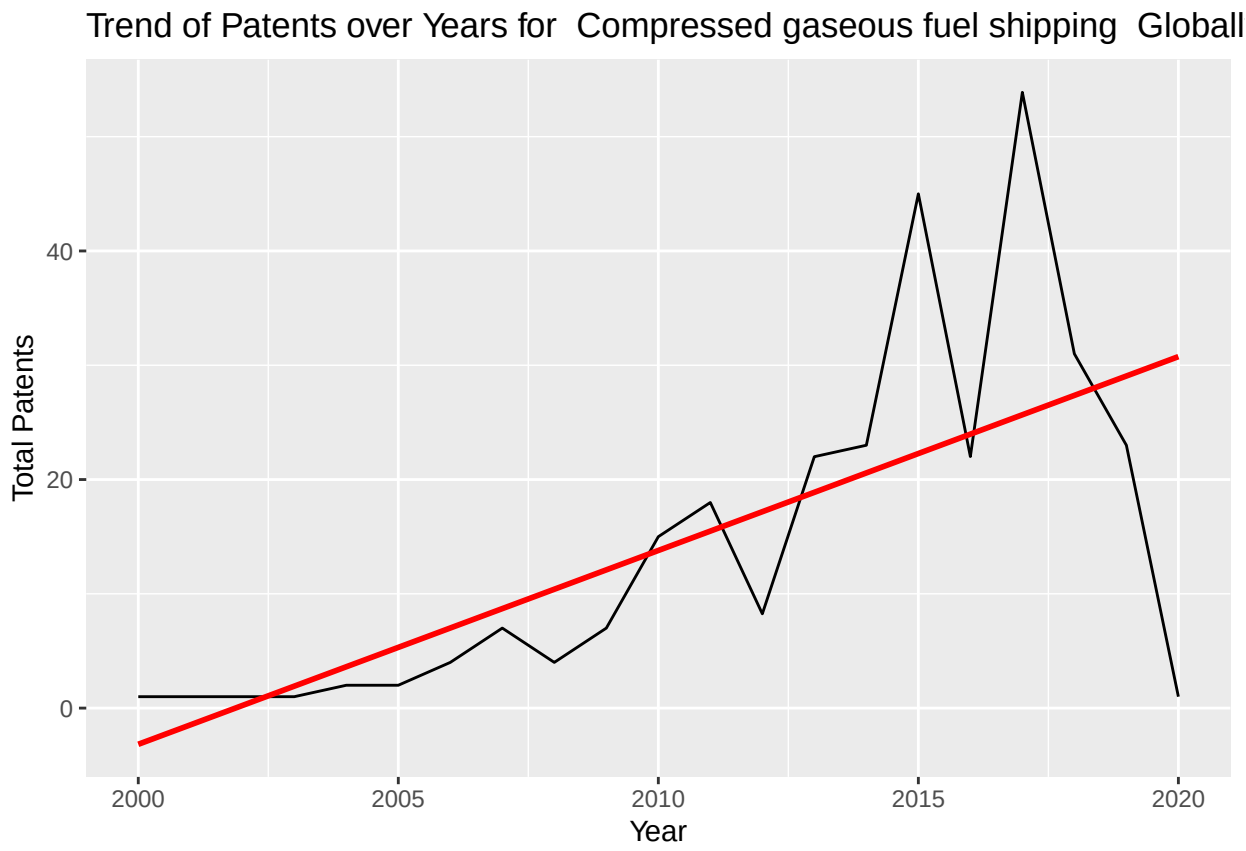

```
## `geom_smooth()` using formula = 'y ~ x'
```



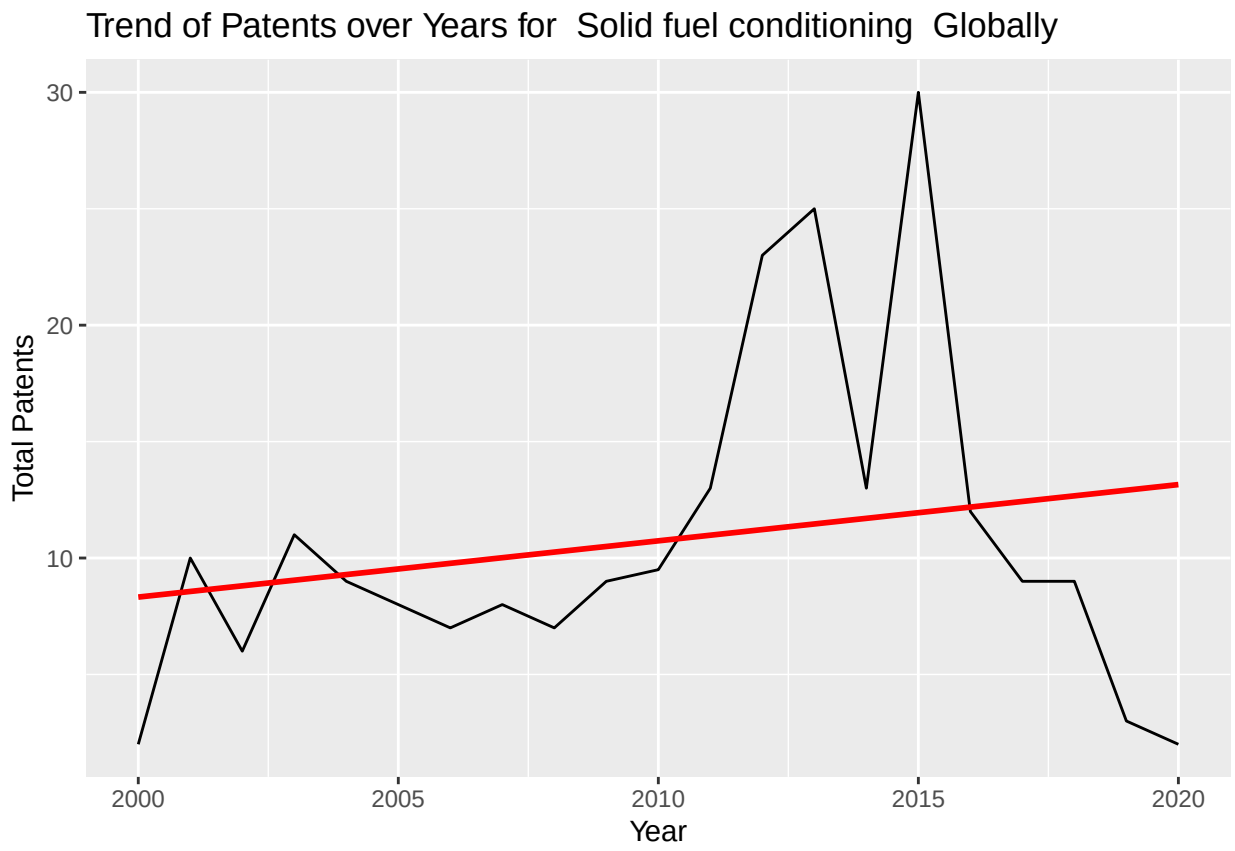

```
## `geom_smooth()` using formula = 'y ~ x'
```



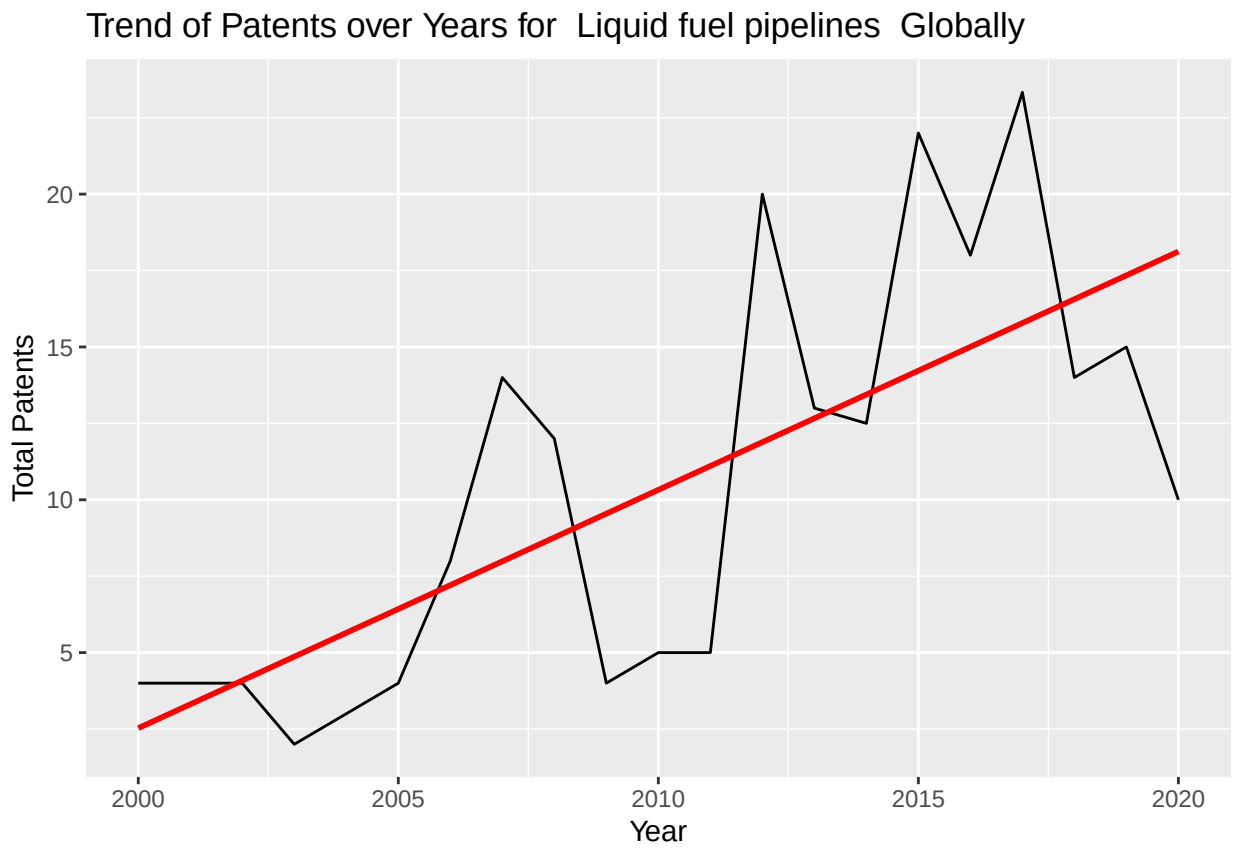

```
## `geom_smooth()` using formula = 'y ~ x'
```



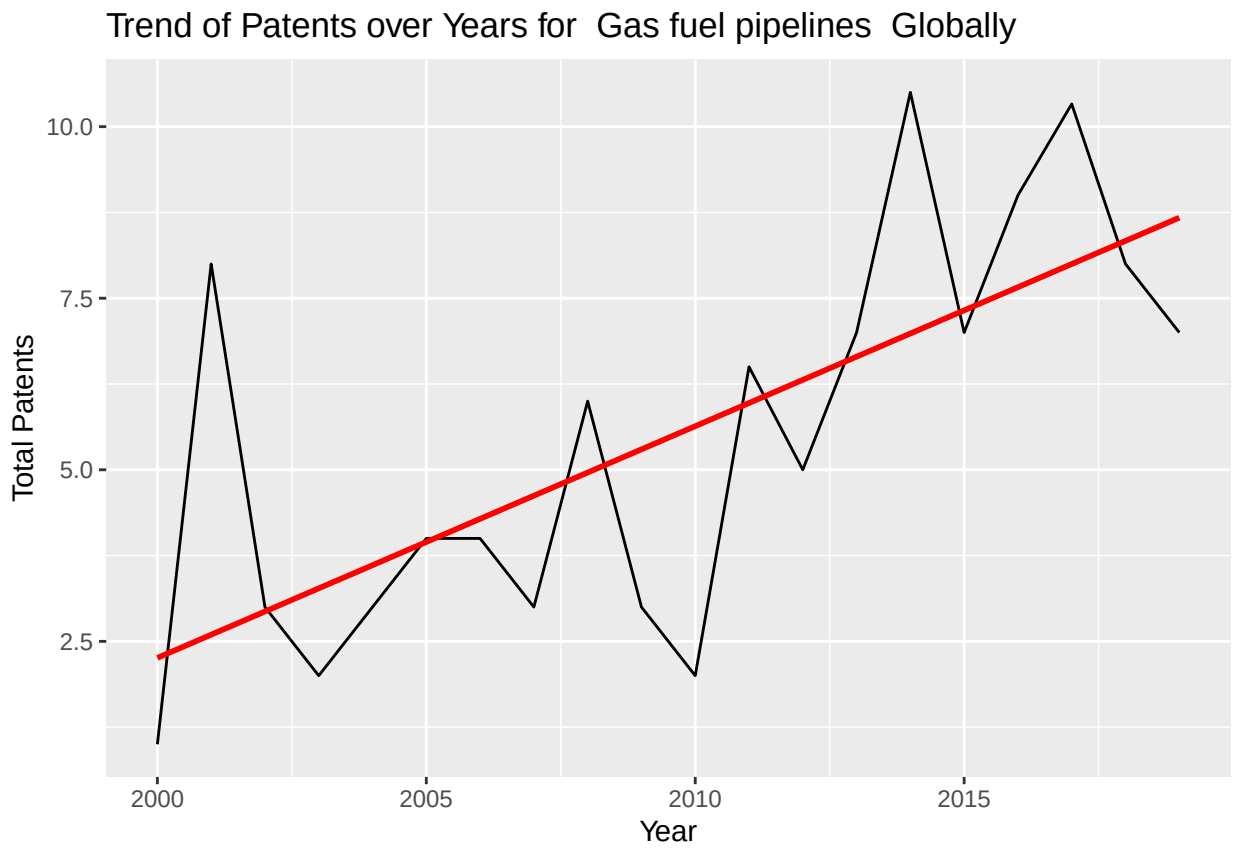

```
## `geom_smooth()` using formula = 'y ~ x'
```



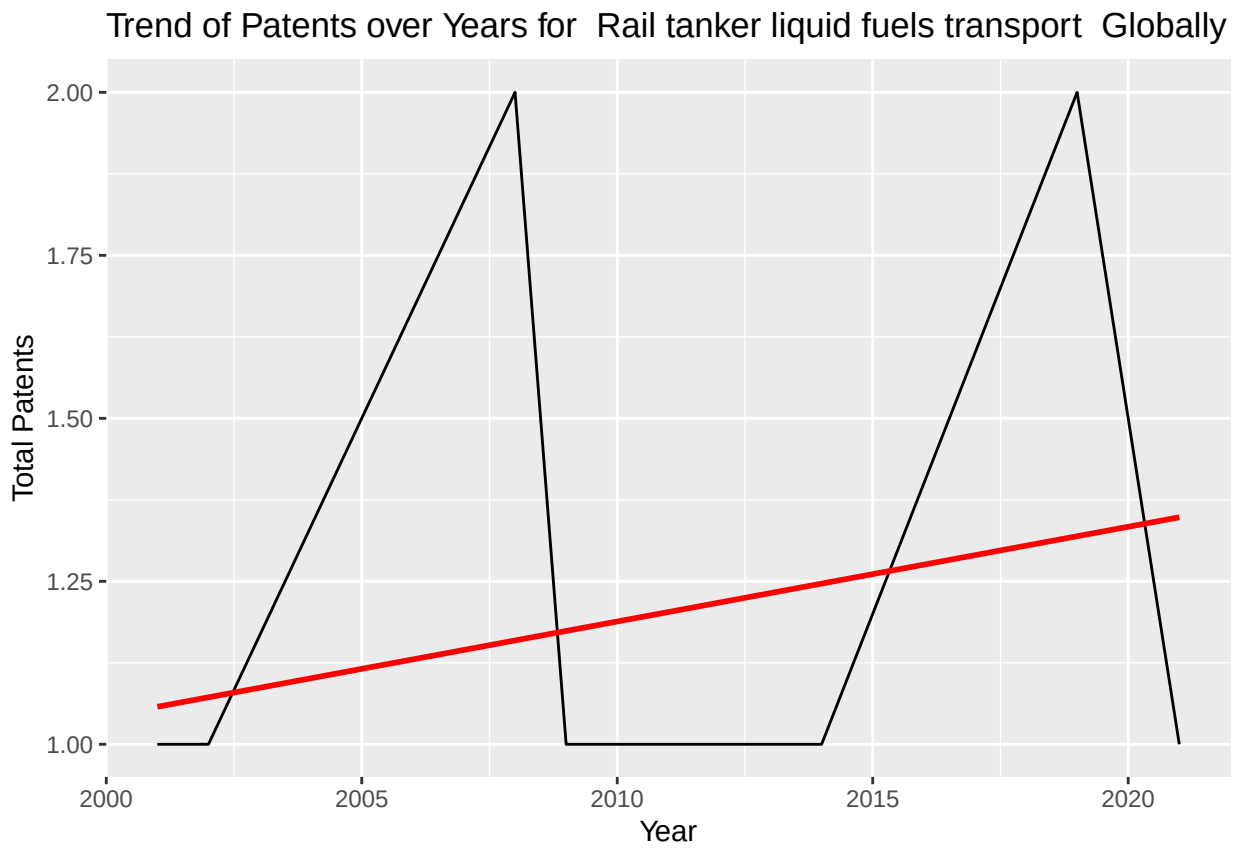

```
## `geom_smooth()` using formula = 'y ~ x'
```



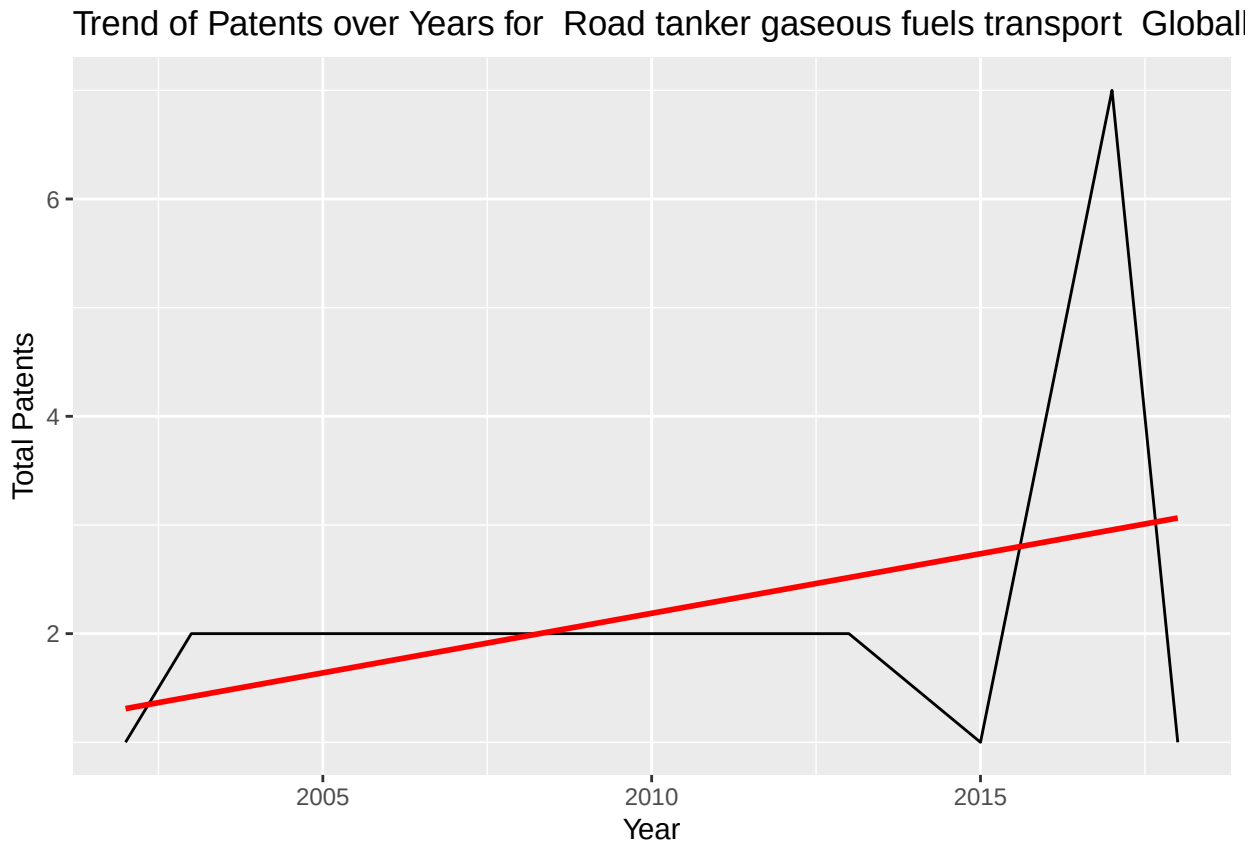

```
## `geom_smooth()` using formula = 'y ~ x'
```



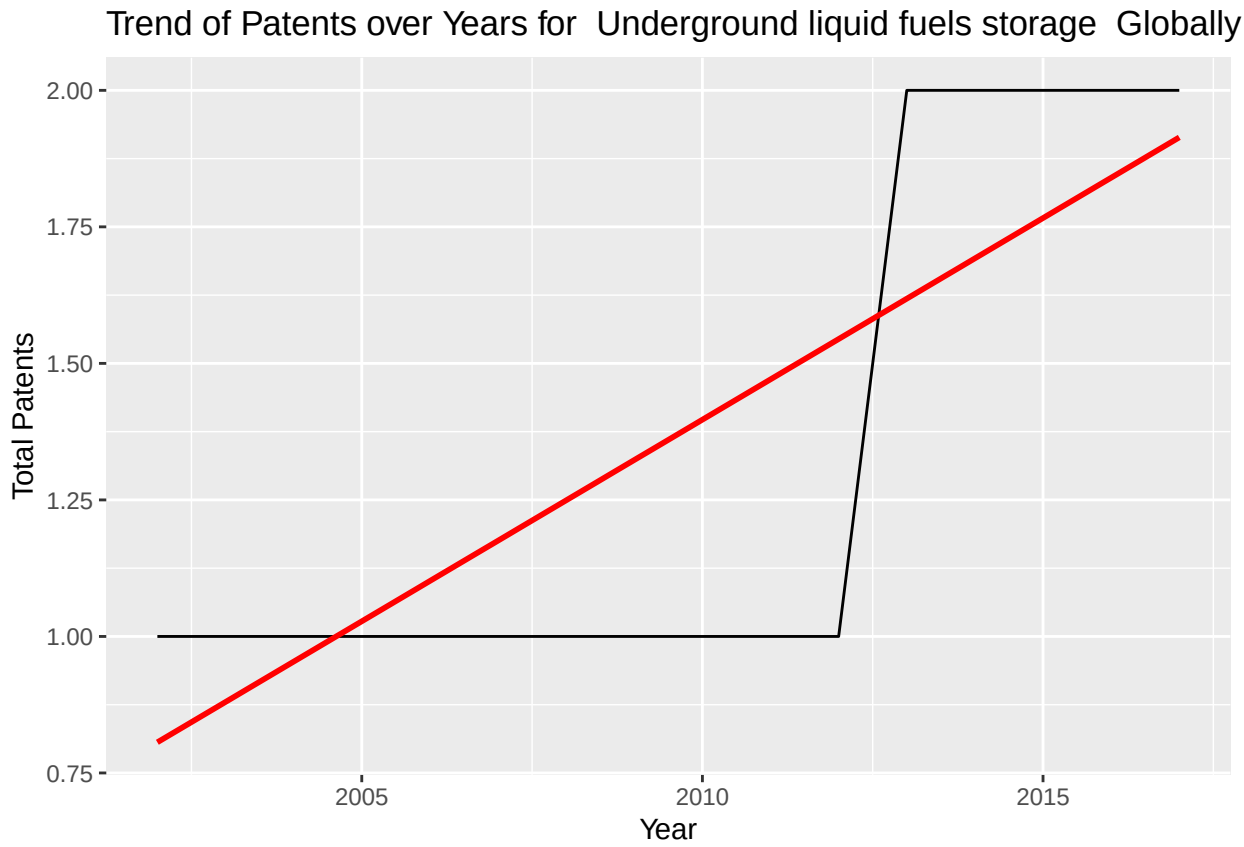

```
## `geom_smooth()` using formula = 'y ~ x'
```



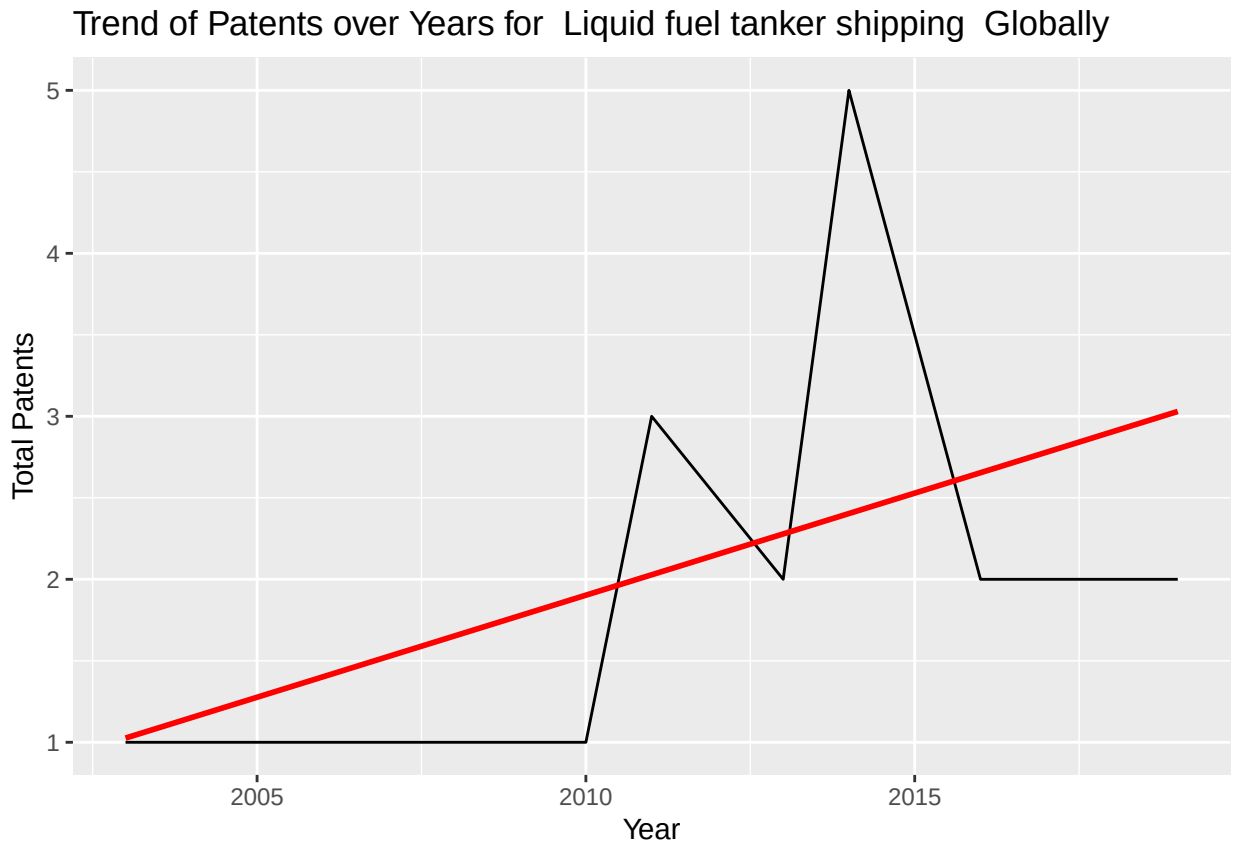

```
## `geom_smooth()` using formula = 'y ~ x'
```



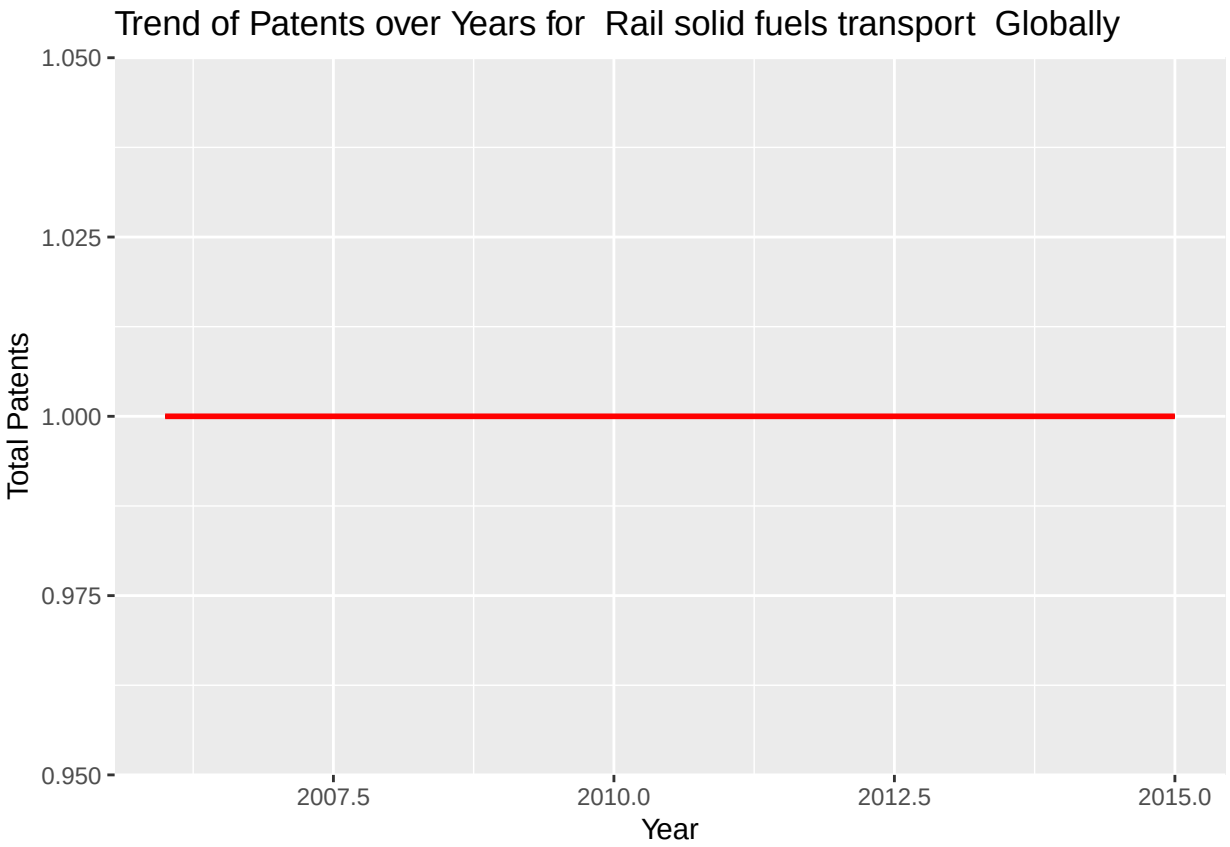

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
## `geom_smooth()` using formula = 'y ~ x'
```




```
## `geom_smooth()` using formula = 'y ~ x'
```


Trend of Patents over Years for Rail solid fuels transport Globally
Total Patents
1.050
1.025
1.000
0.975
0.950
2007.5
2010.0
2012.5
2015.0
Year