```
In [18]: import pandas as pd
         from scipy.stats import zscore
         import matplotlib.pyplot as plt

         # Step 1: Load the dataset
         file_path = r"C:\Users\exam\Downloads\archive (2)\supermarket_sales - Sheet1.csv"
         data = pd.read_csv(file_path)

         # Step 2: Combine 'Date' and 'Time' into a single datetime column and set as index
         data['Datetime'] = pd.to_datetime(data['Date'] + ' ' + data['Time'])
         data.set_index('Datetime', inplace=True)
         data.drop(['Date', 'Time'], axis=1, inplace=True)

         # Step 3: Select numerical columns only and resample to daily frequency
         numerical_columns = data.select_dtypes(include=['float64', 'int64']).columns
         data_daily = data[numerical_columns].resample('D').sum()

         # Step 4: Detect and remove outliers based on the z-score (e.g., 'Total' column)
         data_daily['z_score_total'] = zscore(data_daily['Total'], nan_policy='omit')
         data_cleaned = data_daily[abs(data_daily['z_score_total']) <= 3].drop(columns=['z_score_total'])

         # Visualization: Plot the 'Total' column over time
         plt.figure(figsize=(12, 6))
         plt.plot(data_cleaned.index, data_cleaned['Total'], color='blue', label='Total Sales')
         plt.title('Total Sales Over Time (Cleaned Data)', fontsize=16)
         plt.xlabel('Date', fontsize=12)
         plt.ylabel('Total Sales', fontsize=12)
         plt.grid(True)
         plt.legend()
         plt.show()
```


Total Sales Over Time (Cleaned Data)