**fĐẠI HỌC QUỐC GIA TP HCM**



TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

🙡🕮🙣



**LAB 5**

Môn học: Phân tích dữ liệu kinh doanh

Giảng viên hướng dẫn: **PGS.TS. Nguyễn Đình Thuân**

Lớp: Phân tích dữ liệu kinh doanh – IS403.K22.HTCL

Nhóm 3

Thành viên:

|  |  |  |
| --- | --- | --- |
| Nguyễn Minh Nhựt | - | 17520867 |
| Hoàng Thụy Trinh | - | 17521162 |
| Trần Quang Phúc | - | 17520912 |
| Nguyễn Thị Nhật Hằng | - | 17520437 |
| Lê Ngọc Long | - | 17520710 |

**TP Hồ Chí Minh tháng 05 năm 2020**

**MỤC LỤC**

[NHẬN XÉT CỦA GIÁO VIÊN 1](#_Toc41422290)

[CHƯƠNG 1. ĐẶT VẤN ĐỀ 2](#_Toc41422291)

[CHƯƠNG 2. THỰC HIỆN BÀI TẬP 3](#_Toc41422292)

[1. BÀI TẬP 4A – Home Market Value 3](#_Toc41422293)

[1.1. Thực hiện trên công cụ SPSS 3](#_Toc41422294)

[1.2. Thực hiện trên ngôn ngữ R 7](#_Toc41422295)

[2. BÀI TẬP 4A – Colleges and Universities 9](#_Toc41422296)

[2.1. Thực hiện trên công cụ SPSS 9](#_Toc41422297)

[2.2. Thực hiện trên ngôn ngữ R 13](#_Toc41422298)

[3. BÀI TẬP 4B – Dữ liệu thực tế Việt Nam 17](#_Toc41422299)

[3.1. Thực hiện trên công cụ SPSS 17](#_Toc41422300)

[3.2. Thực hiện trên ngôn ngữ R 20](#_Toc41422301)

[CHƯƠNG 3. PHÂN CÔNG CÔNG VIỆC 24](#_Toc41422302)

[CHƯƠNG 4. TÀI LIỆU THAM KHẢO 26](#_Toc41422303)

# **NHẬN XÉT CỦA GIÁO VIÊN**

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………………

# **CHƯƠNG 1. ĐẶT VẤN ĐỀ**

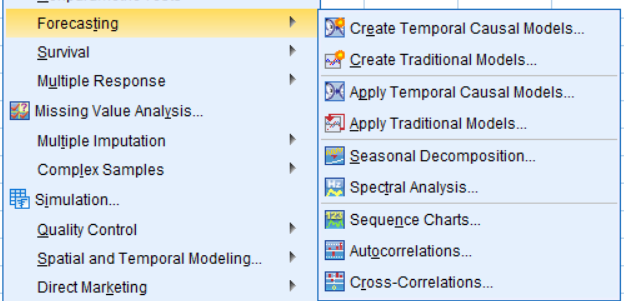
Đối với một doanh nghiệp, việc kinh doanh muốn đạt hiệu quả, tăng trưởng doanh, doanh số, lợi nhuận, … thì cần thiết phải có phân tích dữ liệu kinh doanh (Business Analytics) trong doanh nghiệp. Phân tích dữ liệu kinh doanh góp phần quan trọng để giúp doanh nghiệp đưa ra quyết định đúng đắn như trong trường hợp doanh nghiệp muốn biết số lượng giá các sản phẩm sẽ tăng hay giảm trong các tháng tới để đưa ra quyết định tăng hoặc giảm nguồn cung cấp....

Nội dung bài thực hành số 5: Thực hành dùng ngôn ngữ R và SPSS thực hiện về các phép tính dự báo với các tập dữ liệu: *Tablet Computer Sales, Coal Production, Gas & Electric, Dữ liệu Việt Nam tùy chọn.*

# **CHƯƠNG 2. THỰC HIỆN BÀI TẬP**

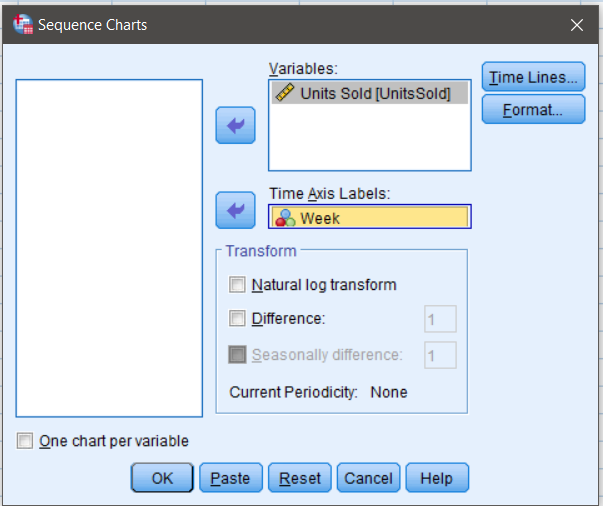
1. **BÀI TẬP 5A – Tablet Computer Sales**
   1. **Thực hiện trên ngôn ngữ R**
   2. **Thực hiện trên công cụ SPSS**

* Xác định dạng đồ thị của Unit Sold theo thời gian (Week) vào công cụ Analyze 🡪 chọn tiếp Forcasting 🡪 Chọn Sequence Chart



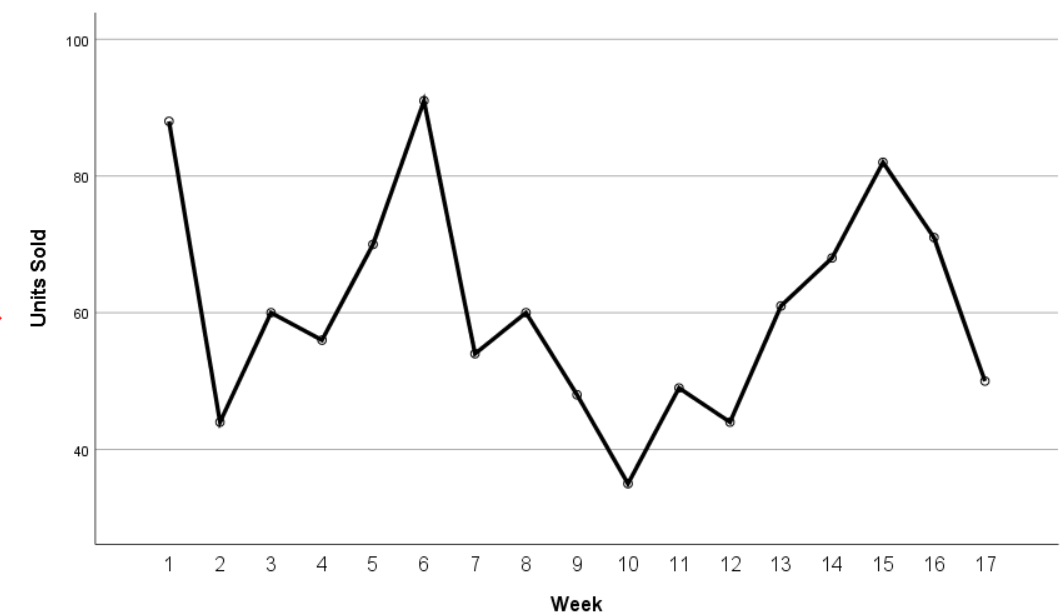
*Hình 5.1 Màn hình chọn chức năng vẽ biểu đồ dự báo theo chuỗi thời gian*

* Chọn nhãn thời gian là Week và Giá trị là Unit Sold nhấn OK

**

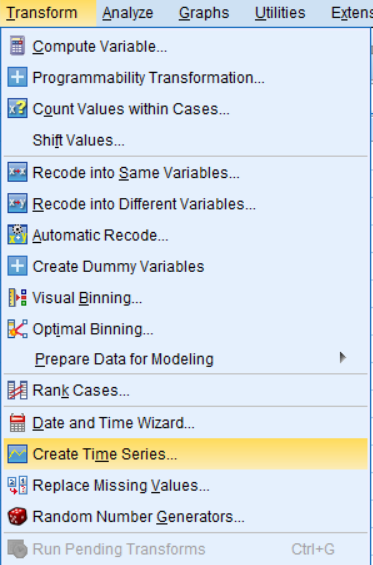
*Hình 5.2 Màn hình chọn của Sequence Charts*

* Ta được đồ thị như sau:

**

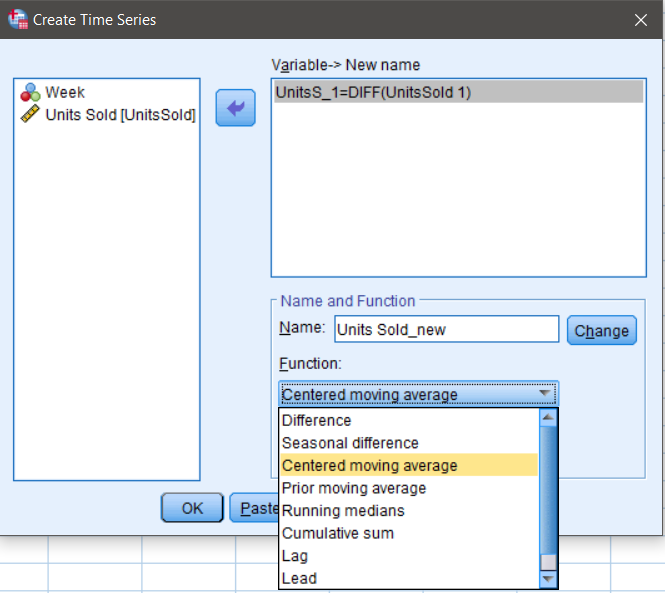
*Hình 5.3 Đồ thị dự báo Units Sold theo Week*

* Tiến hành Smoothing bằng phương pháp **Simple Moving Average**
  + Bước 1: Ta chọn Transform 🡪 Create Time Series

**

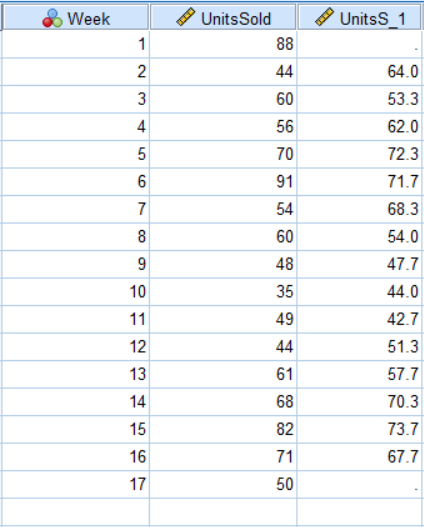
*Hình 5.4 Màn hình chọn chức năng tạo chuỗi thời gian*

* + Bước 2: Chọn Variable Name 🡪 New name là **Units Sold,** đặt lại tên, chọn Function **Centered moving average**, chọn span = 3 và sau đó nhấn OK.

**

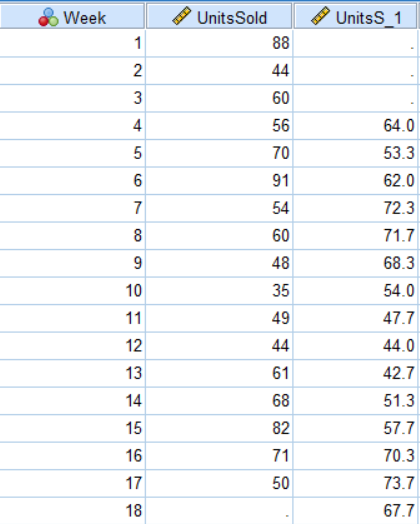
*Hình 5.5 Màn hình Create chuỗi thời gian*

* + Bước 3: Cột UnitsS\_1 sẽ xuất hiện trong tập dữ liệu

**

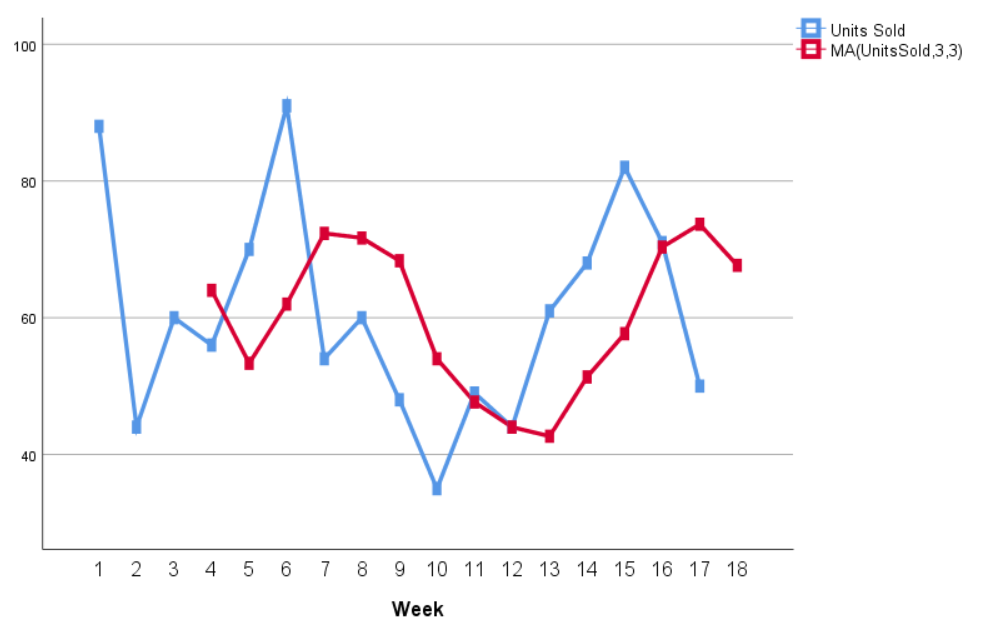
*Hình 5.6 Màn hình với cột UnitsS\_1 đã được thêm*

* + Bước 4: Chỉnh sửa cho tập dữ liệu phù hợp, chỉnh sửa như hình dưới đây

**

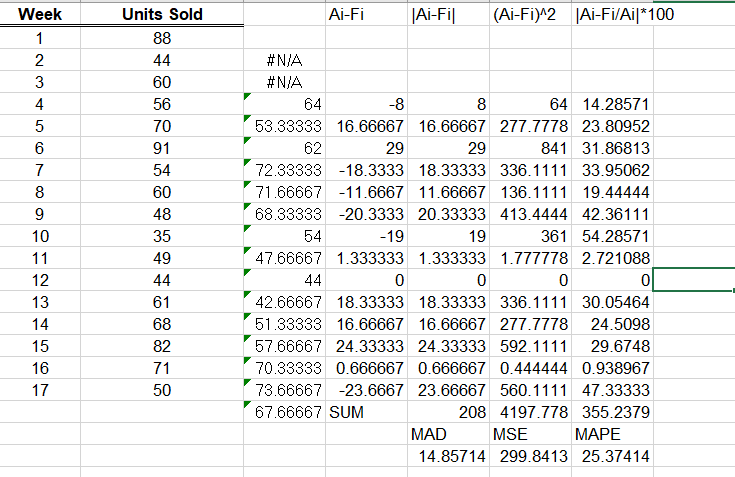
*Hình 5.7 Màn hình dữ liệu đã được chỉnh sửa cho phù hợp*

* + Bước 5: Tiến hành vẽ biểu đồ Sequence để xem kết quả giữa đồ thị ban đầu và đồ thị đã smoothing bằng phương pháp Simple Moving Average.

**

*Hình 5.8 Đồ thị biểu diễn Unit Solid trước khi Smoothing (màu xanh) và sau khi Smoothing (màu đỏ)*

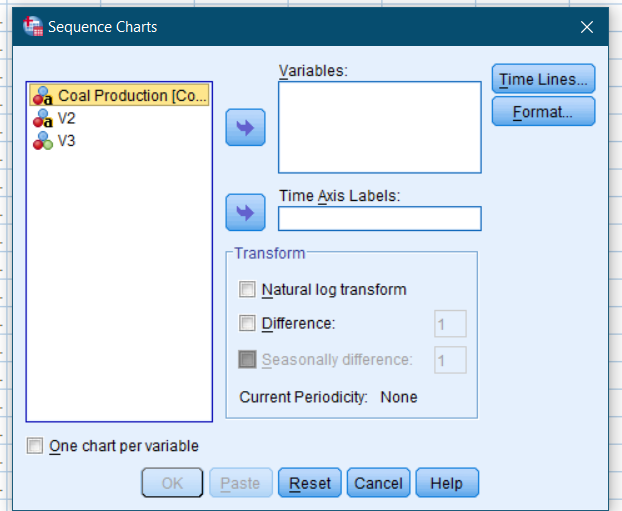
* + Bước 6: Tính các giá trị MAD, MSE, RMSE, MAPE bằng công cụ Excel

**

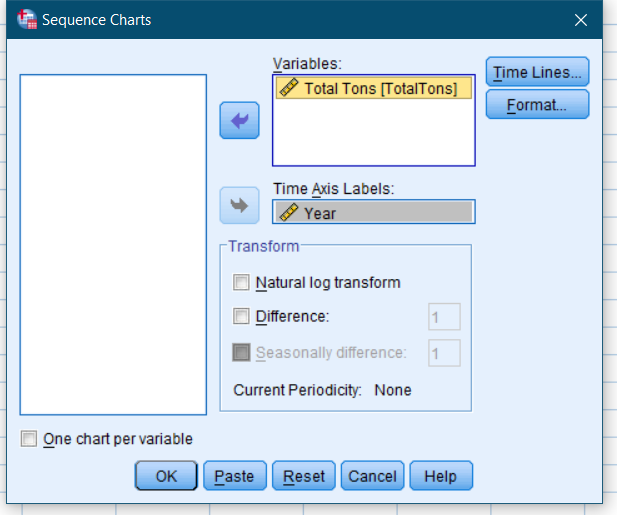
*Hình 5.9 Tính các độ đo đánh giá*

1. **BÀI TẬP 5A – Coal Production**

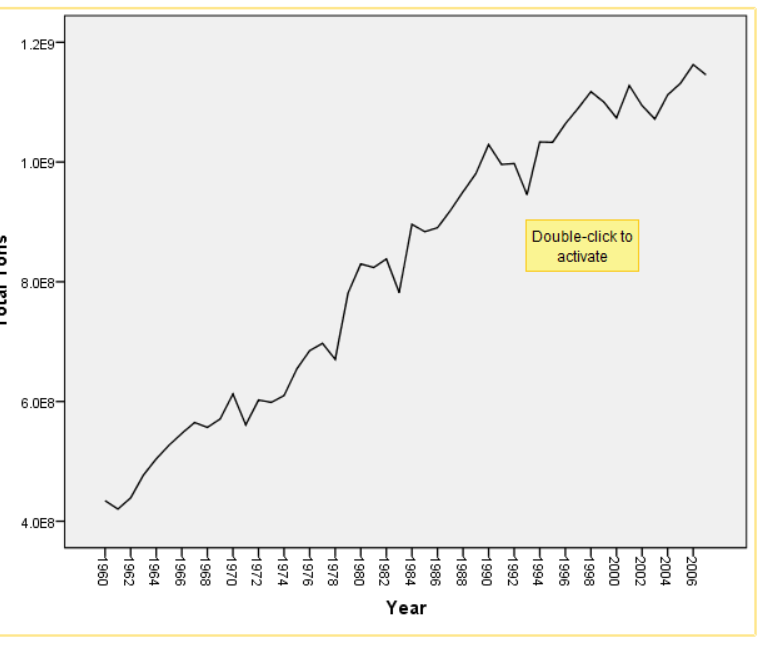
Xác định dạng đồ thị của Unit Sold theo thời gian (Week) vào công cụ Analyze 🡪 chọn tiếp Forcasting 🡪 Chọn Sequence Chart

****

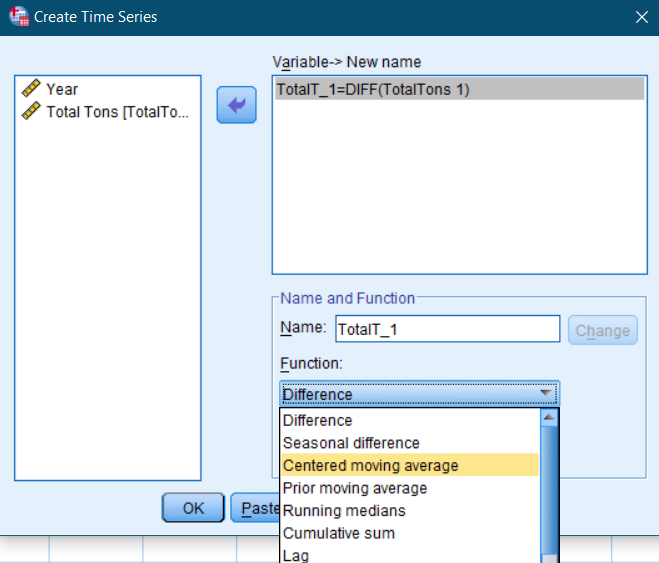
* Chọn nhãn thời gian là Year và Giá trị là Total Tons nhấn OK

****

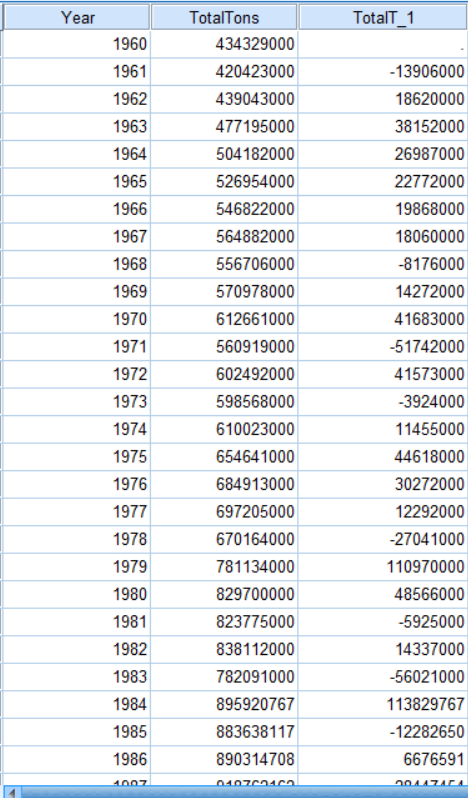
* Ta được đồ thị như sau:

****

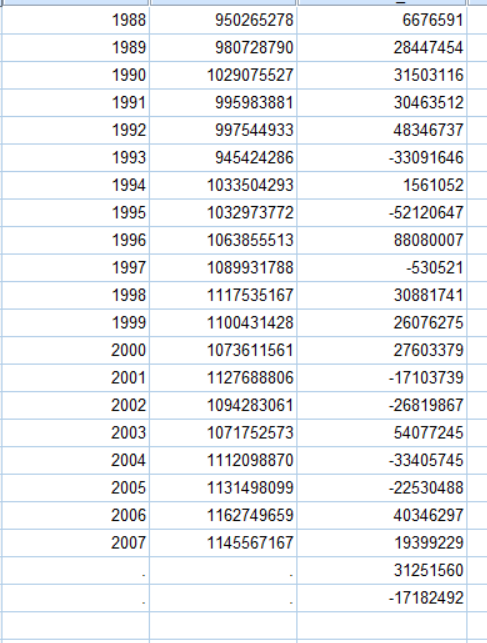
* Tiến hành Smoothing bằng phương pháp **Simple Moving Average**
  + Bước 1: Ta chọn Transform 🡪 Create Time Series

****

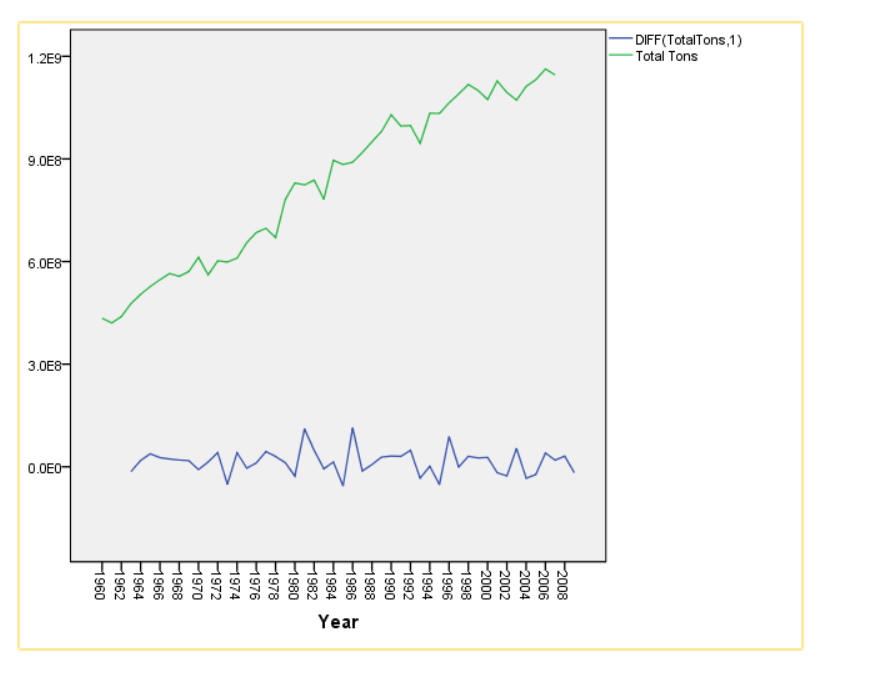
* + Bước 3: Cột TotalT\_1 sẽ xuất hiện trong tập dữ liệu

****

* + Bước 4: Chỉnh sửa cho tập dữ liệu phù hợp, chỉnh sửa như hình dưới đây

****

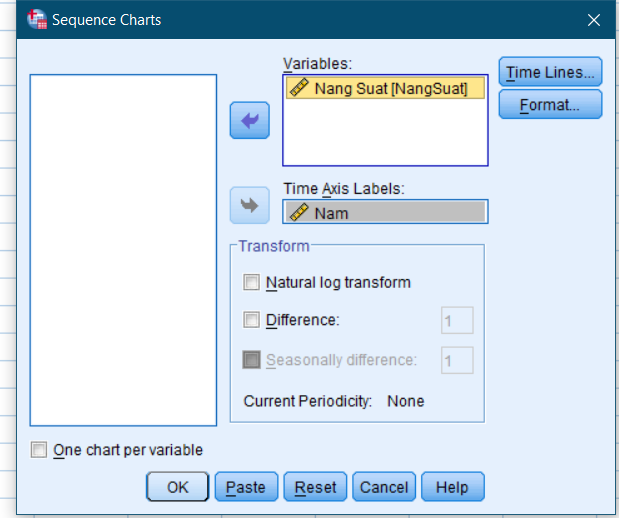
* + Bước 5: Tiến hành vẽ biểu đồ Sequence để xem kết quả giữa đồ thị ban đầu và đồ thị đã smoothing bằng phương pháp Simple Moving Average.

****

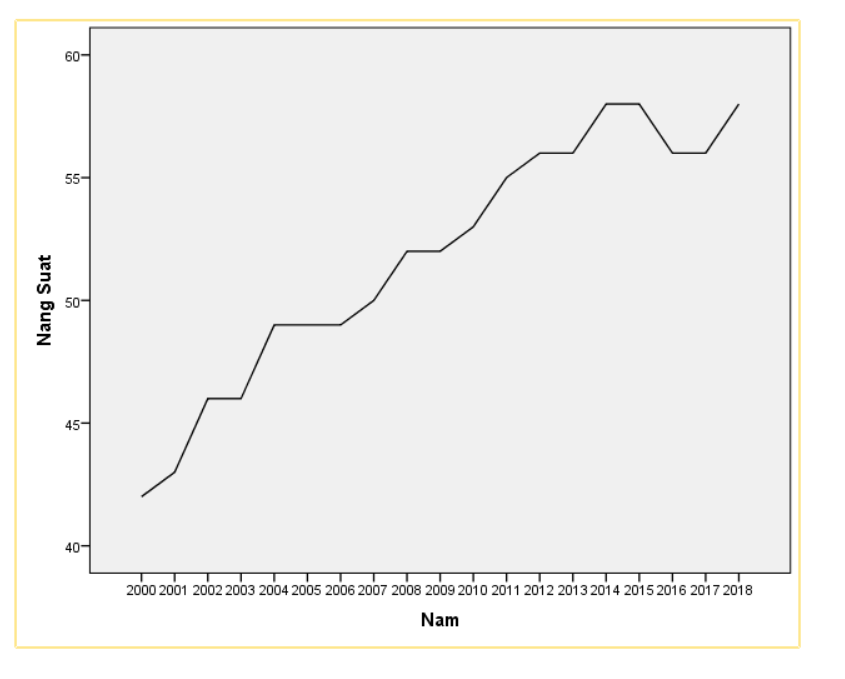
1. **BÀI TẬP 5A – Gas & Electric**
2. **BÀI TẬP 5B – Dữ liệu Việt Nam**

Xác định dạng đồ thị của Năng Suất theo thời gian (Năm) vào công cụ Analyze 🡪 chọn tiếp Forcasting 🡪 Chọn Sequence Chart

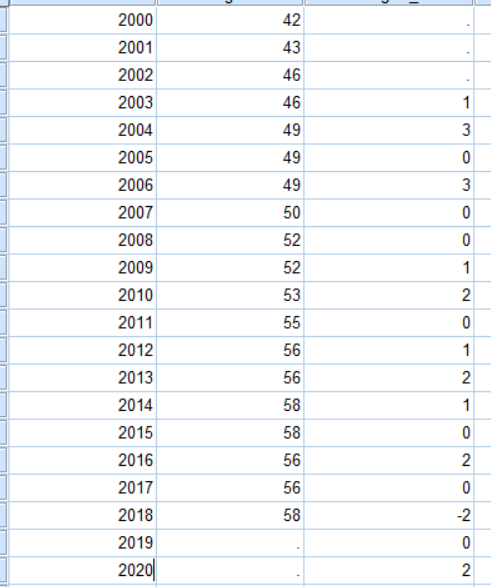
* Chọn nhãn thời gian là Năm và Giá trị là Năng Suất nhấn OK

****

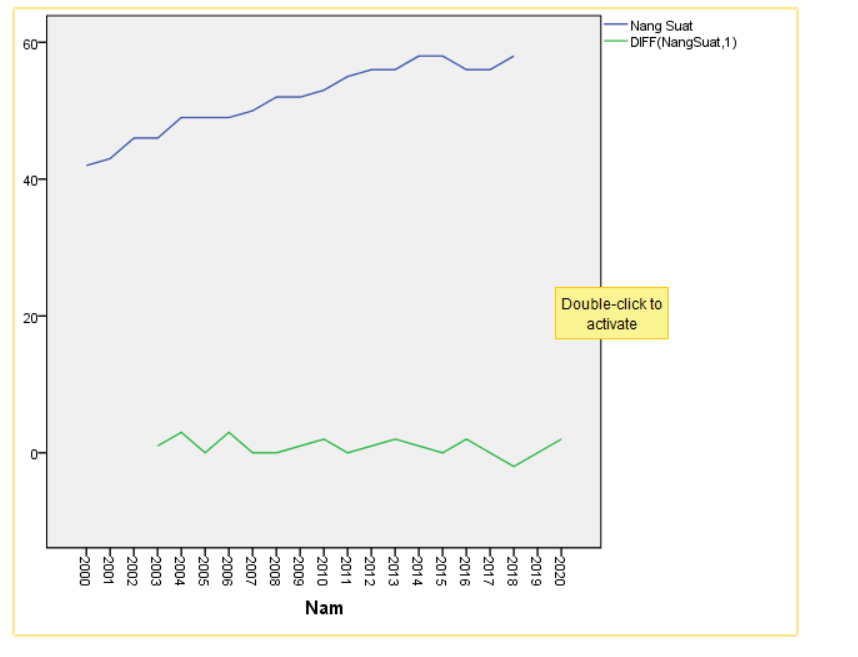
* Ta được đồ thị như sau:

****

* Tiến hành Smoothing bằng phương pháp **Simple Moving Average**
  + Bước 1: Chỉnh sửa cho tập dữ liệu phù hợp, chỉnh sửa như hình dưới đây

****

* + Bước 5: Tiến hành vẽ biểu đồ Sequence để xem kết quả giữa đồ thị ban đầu và đồ thị đã smoothing bằng phương pháp Simple Moving Average.

****

# **CHƯƠNG 3. PHÂN CÔNG CÔNG VIỆC**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Thành viên  Công việc | Nguyễn Minh Nhựt  (Nhóm trưởng) | Hoàng Thụy Trinh | Trần Quang Phúc | Lê Ngọc Long | Nguyễn Thị Nhật Hằng |
| Thực hiện phân tích hồi qui tập dữ liệu Việt Nam trên SPSS | x |  |  |  |  |
| Thực hiện phân tích hồi qui tập dữ liệu Việt Nam trên ngôn ngữ R | x |  |  |  |  |
| Thực hiện phân tích hồi qui tập dữ liệu Home Value trên SPSS |  |  | x |  |  |
| Thực hiện phân tích hồi qui tập dữ liệu Home Value trên R |  | x |  |  |  |
| Thực hiện phân tích hồi qui tập dữ liệu College and Universities trên SPSS |  |  |  |  | x |
| Thực hiện phân tích hồi qui tập dữ liệu College and Universities trên R |  |  |  | x |  |
| Tìm kiếm tài liệu |  | x | x | x | x |
| Hướng dẫn thực hiện | x |  |  |  |  |

# **CHƯƠNG 4. TÀI LIỆU THAM KHẢO**

1. Slide Business Analysis Chapter 9 do giảng viên cung cấp
2. Tài liệu ngôn ngữ R do giảng viên cung cấp
3. Phân tích hồi qui trên công cụ SPSS, Link: <https://www.phamlocblog.com/2016/11/cach-chay-hoi-quy-trong-spss.html>
4. Phân tích hồi qui trên công cụ SPSS, Link: <http://phanmemspss.com/phan-tich-data/regression/hoi-qui-regression-trong-spss.html>
5. Các tiêu chí đánh giá mục 3, Link: [http://rstudio-pubs-static.s3.amazonaws.com/445130\_e065fc3cceaf4393ba8011e3d7e106b5.html](http://rstudio-pubs-static.s3.amazonaws.com/445130_e065fc3cceaf4393ba8011e3d7e106b5.html?fbclid=IwAR32Ev3cWxamrlmrM30AmsMTdqTW-uFK08gaeWNFxKHMLgCO-CzsTDh44Co)