# Project 1 – Trend Analysis

The task of the first „Trend Analysis" Project is presentation of a chosen group of aircrafts, planes, helicopters and gliders built and used over the last few decades as well as analysis of their parameters and forecast for the next 10 years. This project, in a form of presentation, should include the following parts:
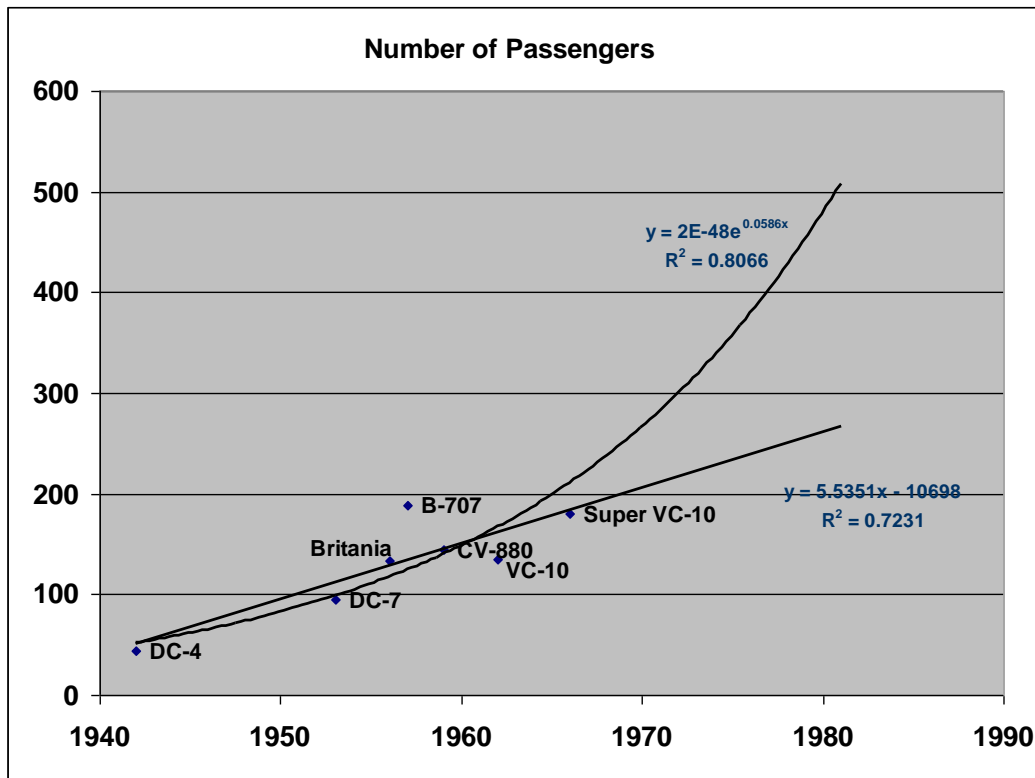
a) definition of the chosen aircraft group,
b) presentation of typical constructions for his group,
c) trend analysis of chosen technical data and performance,
d) summary, including an attempt to pinpoint important features of an aircraft (helicopter, glider), that would be produced in 10 years.

The presentation should take no longer than 10 minutes.

Ad. a) – this part includes specifying characteristic features for the chosen aircraft group; e.g. for a group defined as "piston trainer", the receiver (user) of the plane can be specified, typical requirements can be shown (e.g. the requirement to start from a grassy airfield), etc.

Ad. b) – presentation of typical constructions is a part that requires careful selection of aircrafts from the relevant group; these should be constructions that are typical representatives of the defined group, in the relevant period, produced in major numbers and being nearly a standard in a group (e.g. Embraer Tucano as a representative of the turboprop trainer in the 80s.) Presentation of 2-3 typical constructions for 10-year period is suggested. Analysis should include 30-40 years back.

 Ad. c) – in this part several most important (for the relevant group of aircrafts) technical data (mass, geometrical) and performance parameters should be analysed; they should be presented in relevance to the year, in which the aircraft came into use, or the year of the first flight. The next stage is finding the trend curves. This can be done using proper options of spreadsheets or graph-creating programs, or „manually", using the Least squares method (look below). It is important, to check the trend curve's quality, for example by Pearson's correlation coefficient (R-square factor in MS Excel). However, you should remember that it isn't the only correlation coefficient, and that basing on just one factor, with slightly wrong data, wrong conclusions can be drawn. An Example on Fig.1 shows, that in spite of adjusting the exponential curve Belter (higher $R^2$), the trend line can lack sense – according to it, the number of passengers in a typical "transatlantic" plane in the year 1980 should overcome 500, and reach nearly 900 ten years later.

**Number of Passengers**

$y = 2E\text{-}48e^{0.0586x}$
$R^2 = 0.8066$

$y = 5.5351x - 10698$
$R^2 = 0.7231$

B-707

Super VC-10

Britania    CV-880
            VC-10

DC-7

DC-4

Rys.1 – Example of trend analysis with right (linear) and wrong (exponential) assumed function of regression

Ad. d) – on the basis of trend analysis results, forecast for 10 next years should be done, allowing assumptions for a newly designed plane to be created; the most important mass parameters should be given, as well as geometrical data and assumed performance for the new construction.

## Data analysis basics

In this chapter basic relations connected with data analysis will be presented, in particular method of assigning function of regression (trend line). Let's consider gathered data as a set of numbers pairs: (argument $x_i$, value $Y_i$). Let's assume that:

$$Y_i = f(x_i) + \varepsilon_i \tag{1}$$

where:

$f(x_i)$ – unknown function,
$\varepsilon_i$ – random mistake with expected value $E(\varepsilon_i) = 0$.

Let's take a functional:

$$J(f) = \sum_{i=1}^{n} (Y_i - f(x_i))^2 \tag{2}$$

Function, that minimalises the value of $J(f)$ is called the estimator of the least squares of unknown regression function $f$.

One of the indicators defining the estimation (approximation) of experiment data by function $f$ quality is factor of Pearson's correlation coefficient, that can be defined as:

$$R = \frac{\sum_{i=1}^{n}(f(x_i) - \overline{f(x)})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(f(x_i) - \overline{f(x)})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \tag{3}$$

where:

$$\overline{f(x)} = \frac{1}{n}\sum_{i=1}^{n} f(x_i) \quad , \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{4}$$

The higher the factor's value, the higher is the linear relation between experimental and approximated values. $R = 0$ means no linear relation, $R = 1$ means precise positive linear relation. „R-square" (in MS Excel program) is in fact square of the factor defined above.

## Linear regression

If we assume, that an unknown function $f$ has the form of a function describing a straight line:

$$f(x) = ax + b \tag{5}$$

then the factors of this function that satisfy the condition of minimum of the functional defined in formula (2) can be appointed from the relation:

$$a = \frac{n\sum_{i=1}^{n} x_i Y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} x_i^{2} - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{6}$$

$$b = \frac{\sum_{i=1}^{n} x_i^{2} \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i Y_i}{n\sum_{i=1}^{n} x_i^{2} - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{7}$$

Referring above formulas to the example show on Fig.1, variable $x$ is time (year), and $Y$ is the value of examined parameter(number of passengers).

## References:

1. Siegmund Brandt: Data Analysis: Statistical and Computational Methods for Scientists and Engineers, Springer, 1999
2. Daniel P. Raymer: Aircraft Design: A Conceptual Approach, AIAA Education Series 2012
3. John D. Anderson, Jr.: Introduction to Flight, McGraw Hill, 2005