Gage R&R

May 22, 2012

Andrea Spanò
andrea.spano@quantide.com

- ISO 9000 is a family of **standards related to quality management systems** and designed to help organisations to ensure that they meet the needs of customers and other stakeholders.

- Up to the end of December 2009, at least 1'064'785 ISO 9001 (2000 and 2008) certificates have been issued in 178 countries and economies. (ISO Survey 2009)

- In Italy, 13'066 certificates have been issued. Italy is the European leader and among the firsts in the world for number of ISO 9001 certificates. (ISO Survey 2009)

⋆ Some definitions:

- **quality**: the degree to which a set of inherent **characteristics** fulfils **requirement**;
- **management**: coordinated activities aimed to direct and control;
- **quality management system**: the organizational structure, procedures, processes and resources aimed to **direct** and **control** an **organisation** with regard to **quality**.

⋆ **Monitoring and Measurement of Product**: The organisation shall monitor and **measure the characteristics of the product** to verify that **product requirements** have been met.

⋆ **ISO 9001** describes **Gage R&R** as one of the quality tools to use in support of production/process controls, verification/validation activities, as well as calibration programs.

- In any research or production environment, **measurements system requires to be validated** prior its usage. Any measurement system must be "adequate" to detect differences between parts. The more the quality of a process increases, the more the measurement systems are required to be sophisticated in order to be able to measure the process.

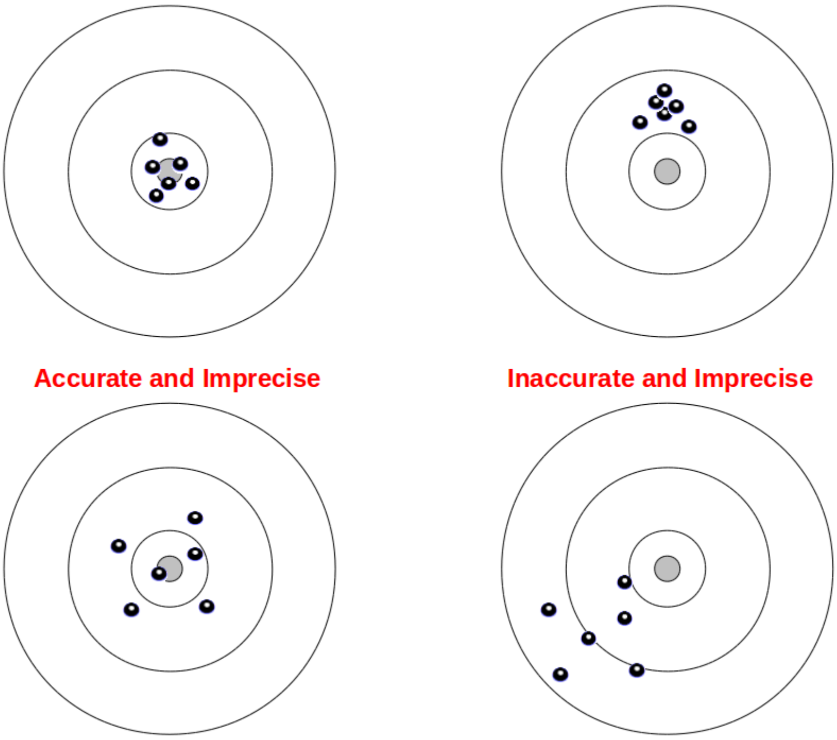
- The ANOVA Gage R&R statistical method estimates **the amount of variability induced in measurements by the measurement system itself**, and compares it to the total variability observed in order to determine the viability of the measurement system.

- This seminar is about **understanding the components and the statistics** beyond an ANOVA based Gage R&R studies for non-destructive tests.

- **Gage R&R : Gage Repeatability and Reproducibility**

- **Repeatability**: it is the ability of measurement system to obtain small variability by repeating the same measurements on the same sample under the same conditions.
- **Reproducibility**: it is the ability of measurement system to return consistent measurements while varying the measurement conditions (different operators, different humidity, etc.)
- **Accuracy**: it is the ability of measurement system to return average measurements close the reference sample (when it exists) for which the "real value" is known.

- **Two fundamental aspects** can be defined for a gage: **precision** (repeatability and reproducibility) and **accuracy**.



Aim of Gage R&R is to assess the **precision** of measurement system.

Main factors affecting a measurement system:

- ⋆ **Measuring instruments:** the gauge or instrument itself and all mounting blocks, supports, fixtures, load cells, etc.
- ⋆ **Operators:** the ability of a person to follow the written or verbal instructions to perform the measurements.
- ⋆ **Parts** (what is being measured): some items are easier to be measured than others; this affects the precision of measurement system.
- ⋆ **Specification:** the measurement is reported against a specification or a reference value. The engineering tolerance does not affect the measurement, but is an important factor in evaluating the viability of the measurement system.

Gage R&R tries to estimate and/or relate all these components each other to assess the "goodness" of the measurement system

- ⋆ **Choose the number and the type of samples (Parts) to be used**. They must be chosen so as to cover all the operative process range and the variability within it.
- ⋆ **Measurement conditions**, usually due to operators' alternation (Operators).
- ⋆ **Number of repeated measurements** in different times, after previous points are set. The number of repeated measurements usually range from 2 to 5.
- ⋆ **Randomize measurements order** by varying previous points. This is required to minimise systematic errors.

- **Average & Range** (**A&R**): it can be used when the number of repetitions is two or three. This method uses the range to estimate repeatability and reproducibility of measurement systems.

- **Average & Variance** (**A&V**) or **ANOVA**: this method is more flexible than the previous one because it can be used with any number of samples, operators and repetitions. Also this method is used to estimate repeatability and reproducibility of measurement systems.

The following equation holds:

$$\sigma^2_{total} = \sigma^2_{product} + \sigma^2_{gage}$$

Total variability can be split into two components: a component intrinsic to the product (or process), and a component due to the measurement error, or **GAGE variability**.

Goals of method:

- Assessing the maximum variability of a measurement system by separating the variability due to the measured process from the variability due to the measurement system within the total variation;
- Evaluating the need (usually based on the precision/tolerance ratio) to improve the measurement system and/or instrument.

When:

- ⋆ At start-up phase of a measurement instrument or bench;
- ⋆ To qualify a new operator;
- ⋆ Before using a calibration control chart. This last is actually used to "take a snapshot" about the process, which will be monitored over time;
- ⋆ After the instruments calibration;
- ⋆ Every time one suspects that the measurement system is not enough precise with respect to tolerances.

This is the most recent method (between the two previously listed) used to estimate the variance components due to repeatability and reproducibility. The quantities calculated with ANOVA can be used to obtain the indices previously shown.

The starting point is the "basic" identity:

$$\sigma^2_{Total} = \sigma^2_{Parts} + \sigma^2_{Gage}$$

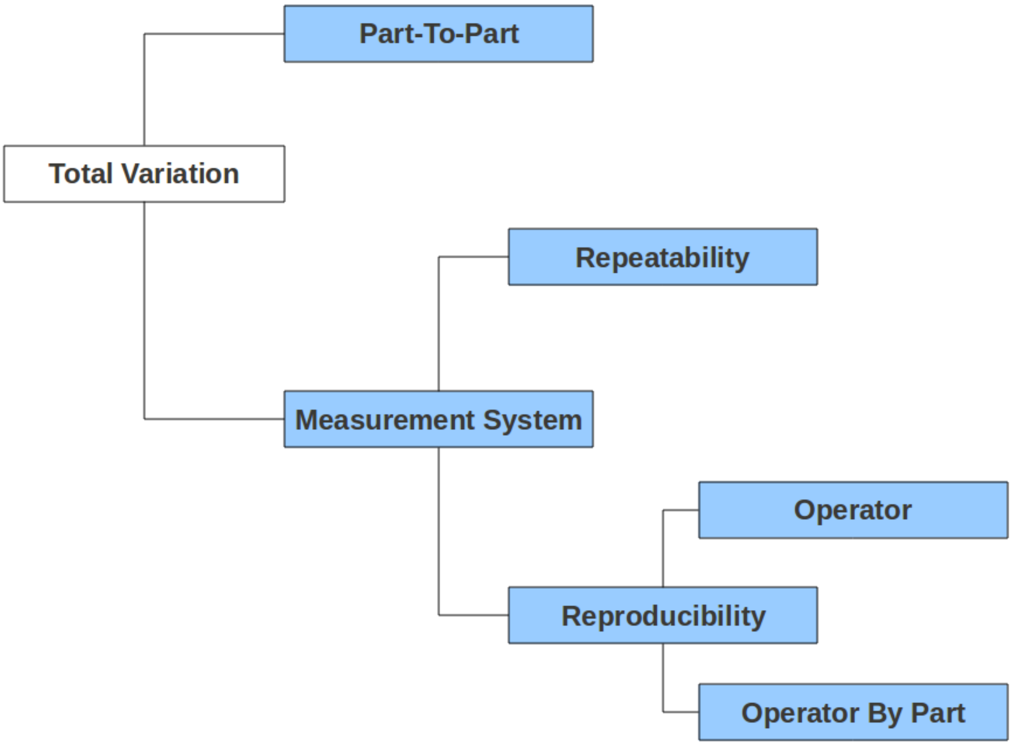
where

$$\sigma^2_{Gage} = \sigma^2_{Reproducibility} + \sigma^2_{Repeatability}$$

and then

$$\sigma^2_{Total} = \sigma^2_{Parts} + \sigma^2_{Reproducibility} + \sigma^2_{Repeatability}$$

Note: $\sigma^2_{Reproducibility}$ is the variability due to change in experimental conditions, i.e., to operators' change, and to their interactions with the measured parts.



Suppose to have conducted a Two-ways ANOVA analysis using the data from an R&R study. The factors are Operator (with *a* levels) and Parts (with *b* levels). *n* measurements are drawn for each crossing between Operator and Parts levels.

The following ANOVA table is obtained:

Source of Variability	Sum of Squares	Degrees of Freedom	Mean of Squares	F statistics
Operator	SS_A	a-1	$MS_A = \frac{SS_A}{a-1}$	$F = \frac{MS_A}{MS_E}$
Part	SS_B	b-1	$MS_B = \frac{SS_B}{b-1}$	$F = \frac{MS_B}{MS_E}$
Interaction (Op x Part)	SS_{AB}	(a-1)(b-1)	$MS_{AB} = \frac{SS_{AB}}{(a-1)(b-1)}$	$F = \frac{MS_{AB}}{MS_E}$
Error	SS_E	ab(n-1)	$MS_E = \frac{SS_E}{ab(n-1)}$	
Total	SS_T	a·b·n-1		

At the beginning, the simplest quantity to estimate is the variance due to repeatability, i.e., the variance due only to the measurement process, with all other factors (operator and part) held fixed. This quantity may be estimated with MS_E. Then:

$$\hat{\sigma}^2_{Repeatability} = MS_E$$

Also, it may be shown that "good" estimates for σ^2_{Parts}, $\sigma^2_{Operator}$ and $\sigma^2_{Operator \times Parts}$ are, respectively:

$$\hat{\sigma}^2_{Parts} = \frac{MS_B - MS_{AB}}{a \cdot n_{rep}}$$

$$\hat{\sigma}^2_{Operator} = \frac{MS_A - MS_{AB}}{b \cdot n_{rep}}$$

$$\hat{\sigma}^2_{Operator \times Parts} = \frac{MS_{AB} - MS_E}{n_{rep}}$$

Where n_{rep} il the number of replications, within each part, operator.

As a consequence, the variance component due to reproducibility is estimated as the sum of estimated variance due to operators and of estimate variance due to operators by part interaction:

$$\hat{\sigma}^2_{Reproducibility} = \hat{\sigma}^2_{Operator} + \hat{\sigma}^2_{Operator \times Parts}$$

Notes:

- ⋆ The ANOVA method, unlike the Range method, allows to estimate the variance component due to Operator by Part interaction. For example, one may assess if some operators over-estimate the measurements on small parts and under-estimate large parts, while others do not. In other words, one may assess if different operators measure differently the same parts.

- ⋆ If the interaction term is not significant (or if $\hat{\sigma}^2_{Operator \times Parts} \leq 0$) then $\hat{\sigma}^2_{Reproducibility} = \hat{\sigma}^2_{Operator}$.
  In this case, the $\sigma^2_{Operator}$ estimate will become:

  $$\hat{\sigma}^2_{Operator} = \frac{MS_A - MS_{E_{pool}}}{b \cdot n_{rep}}$$

  Where $MS_{E_{pool}} = MS_E$ is obtained by setting equal to zero the interaction effect.
- ⋆ In above example, the variance component due to Operators is not significant, while the component due to interaction between Operator and Part is significant. This implies that $\sigma^2_{Reproducibility}$ is greater than 0.

Another parameter to evaluate the measurement system is the *Number of distinct categories*:

$$NoDC = \frac{\hat{\sigma}_{Parts}}{\hat{\sigma}_{Gage}} \cdot \sqrt{2}$$

That index is the "number of non overlapped confidence intervals" that are needed to cover all the product variability. This index may be thought as the number of groups, within the process, that the measurement system is able to distinguish.
Automobile Industry Action Group (AIAG) suggests that, when the Number of distinct categories is less than 2, the measurement system is unable to measure the process. When the number of distinct categories is 2, the measurements can be split in two groups only: "High" and "Low", etc... A Number of distinct categories greater or equal to 5 means an acceptable measurement system.

- ⋆ In a nozzle manufactoring plant a new digital measurement system has been set-up.
- ⋆ Engineers want to assess the performaces of this new gage using Gage R&R.
- ⋆ Nine nozzles have been randomly sampled from the manufactoring process, by considering all the major source of variation in prodction process (machine, time, shift, working team,...).
- ⋆ Two operators measured the diameter of each nozzle two times; the nozzle were presented to operators in random order.
- ⋆ Tecnichal specs for the nozzle require a diameter between 9012 ± 4 microns.

The resulting data have been collected in a dataframe of 36 rows and five columns:

```
> design = gageRRDesign(Operators=2, Parts=9, Measurements=2, randomize=FALSE)
> response(design) =c(9014,9013.8,9011.6,9011.4,9012.4,9012.2,...,9012.3)
> str(design)
Formal class 'gageRR' [package "qualityTools"] with 19 slots
  ..@ X                 :'data.frame': 36 obs. of  5 variables:
  .. ..$ StandardOrder: int [1:36] 1 2 3 4 5 6 7 8 9 10 ...
  .. ..$ RunOrder     : int [1:36] 1 2 3 4 5 6 7 8 9 10 ...
  .. ..$ Operator     : Factor w/ 2 levels "A","B": 1 2 1 2 1 2 1 2 1 2 ...
  .. ..$ Part         : Factor w/ 9 levels "A","B","C","D",..: 1 1 2 2 3 3 4 4 5 5 ...
  .. ..$ Measurement  : num [1:36] 9014 9014 9012 9011 9012 ...
  ..@ ANOVA          :
...
```

To calculate the variance components based on ANOVA method, the following lines of code are used:

```
> gdo = gageRR(design, method = "crossed", sigma = 6, alpha = .25,tolerance=8)
> summary(gdo)
```

With ANOVA tables:

```
   Operators:  2  Parts:  9
Measurements:  2  Total:  36
----------

AnOVa Table -  crossed Design
              Df  Sum Sq Mean Sq F value Pr(>F)
Operator       1    0.04   0.040   3.600 0.0739 .
Part           8   46.15   5.769 519.175 <2e-16 ***
Operator:Part  8    0.06   0.007   0.675 0.7074
Residuals     18    0.20   0.011
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

----------
AnOVa Table Without Interaction -  crossed Design
            Df  Sum Sq  Mean Sq F value Pr(>F)
(a) Operator     1    0.04   0.040     4.0  0.056 .
(b) Part         8   46.15   5.769   576.9 <2e-16 ***
(c) Residuals   26    0.26   0.010
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


And variance components table:

```
    Gage R&R
                      VarComp VarCompContrib  Stdev StudyVar StudyVarContrib P/T Ratio
(1) totalRR           0.01167        0.00804 0.1080    0.648          0.0897    0.0810
(2)  repeatability    0.01000        0.00689 0.1000    0.600          0.0830    0.0750
(3)  reproducibility  0.00167        0.00115 0.0408    0.245          0.0339    0.0306
       Operator       0.00167        0.00115 0.0408    0.245          0.0339    0.0306
       Operator:Part  0.00000        0.00000 0.0000    0.000          0.0000    0.0000
(4) Part to Part      1.43965        0.99196 1.1999    7.199          0.9960    0.8999
    totalVar          1.45132        1.00000 1.2047    7.228          1.0000    0.9035

    ---
     * Contrib equals Contribution in %
     **Number of Distinct Categories (truncated signal-to-noise-ratio) = 15
```

Note that:

$VarComp(1) = VarComp(2) + VarComp(3)$

$VarComp(2) = MeanSq(c)$

$VarComp(3) = \frac{MeanSq(a) - MeanSq(c)}{NParts * n_{rep}}$

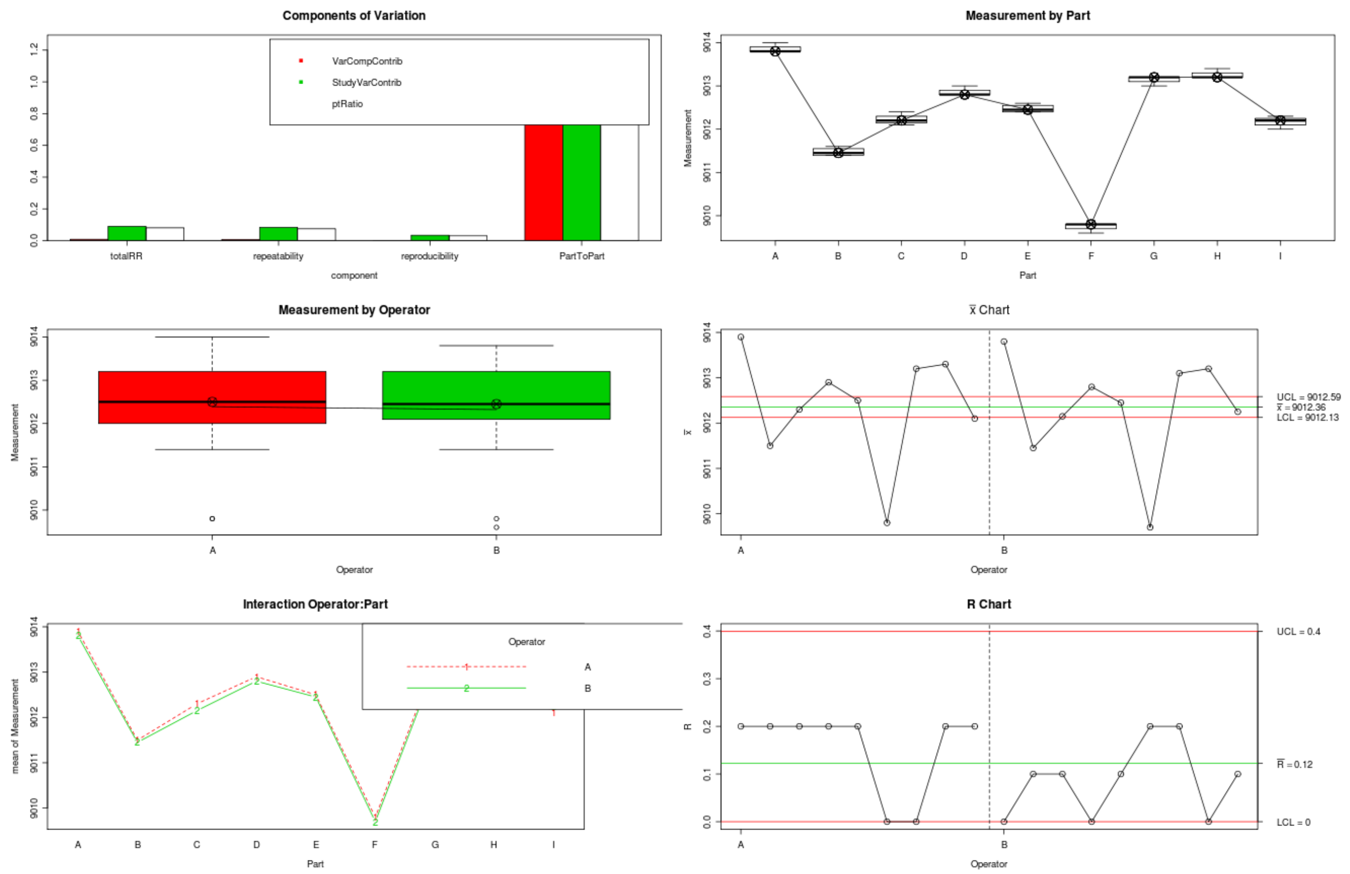
Number of Distinct Categories $= \frac{Stdev(4)}{Stdev(1)} \cdot \sqrt{2}$

Where $NParts$=9 and n_{rep}=2

The following table summarizes the AIAG guidelines for the Gage R&R table

P/T Ratio StudyVarContrib	System is…
Under 10%	Acceptable
10% to 30%	Potentially acceptable (depends on the criticality of the measurement, costs, risks, etc.)
Over 30%	Not acceptable

The next graph summarizes the numerical results of example study



- A company that manufactures BMX wheels had just bought a new measurement instrument that measures the rim strip thickness
- Engineers want to assess the performaces of this new gage using Gage R&R.
- 10 rim strips where randomly selected from the production processes
- Three qualified operators have been selected for the study
- Each operator measured two times each part (rim strip)
- The parts are given to operators in random order
- A total of 60 measures have been drawn
- The tolerance for the rim strip thickness is 0.5 mm.

The resulting data have been collected in a dataframe of 60 rows and five columns:

```
> rimstripGRR = gageRRDesign(Operators=3, Parts=10, Measurements=2, randomize=FALSE)
> response(rimstripGRR) = c(0.22,0.18,0.17,0.33,0.35,0.35,0.28,0.27,0.27, ... ,0.17,0.27)
> str(rimstripGRR)
Formal class 'gageRR' [package "qualityTools"] with 19 slots
  ..@ X                :'data.frame': 60 obs. of  5 variables:
  .. ..$ StandardOrder: int [1:60] 1 2 3 4 5 6 7 8 9 10 ...
  .. ..$ RunOrder     : int [1:60] 1 2 3 4 5 6 7 8 9 10 ...
  .. ..$ Operator     : Factor w/ 3 levels "A","B","C": 1 2 3 1 2 3 1 2 3 1 ...
  .. ..$ Part         : Factor w/ 10 levels "A","B","C","D",..: 1 1 1 2 2 2 3 3 3 4 ...
  .. ..$ Measurement  : num [1:60] 0.22 0.18 0.17 0.33 0.35 0.35 0.28 0.27 0.27 0.28 ...
  ..@ ANOVA            :
...
```

To calculate the variance components base on ANOVA method, the following lines of code are used:

```
> rimstripGRRout = gageRR(rimstripGRR, method = "crossed", sigma = 6, alpha = .25, tolerance = 0.5)
> summary(rimstripGRRout)
```

With following ANOVA table and estimated variation contributions:

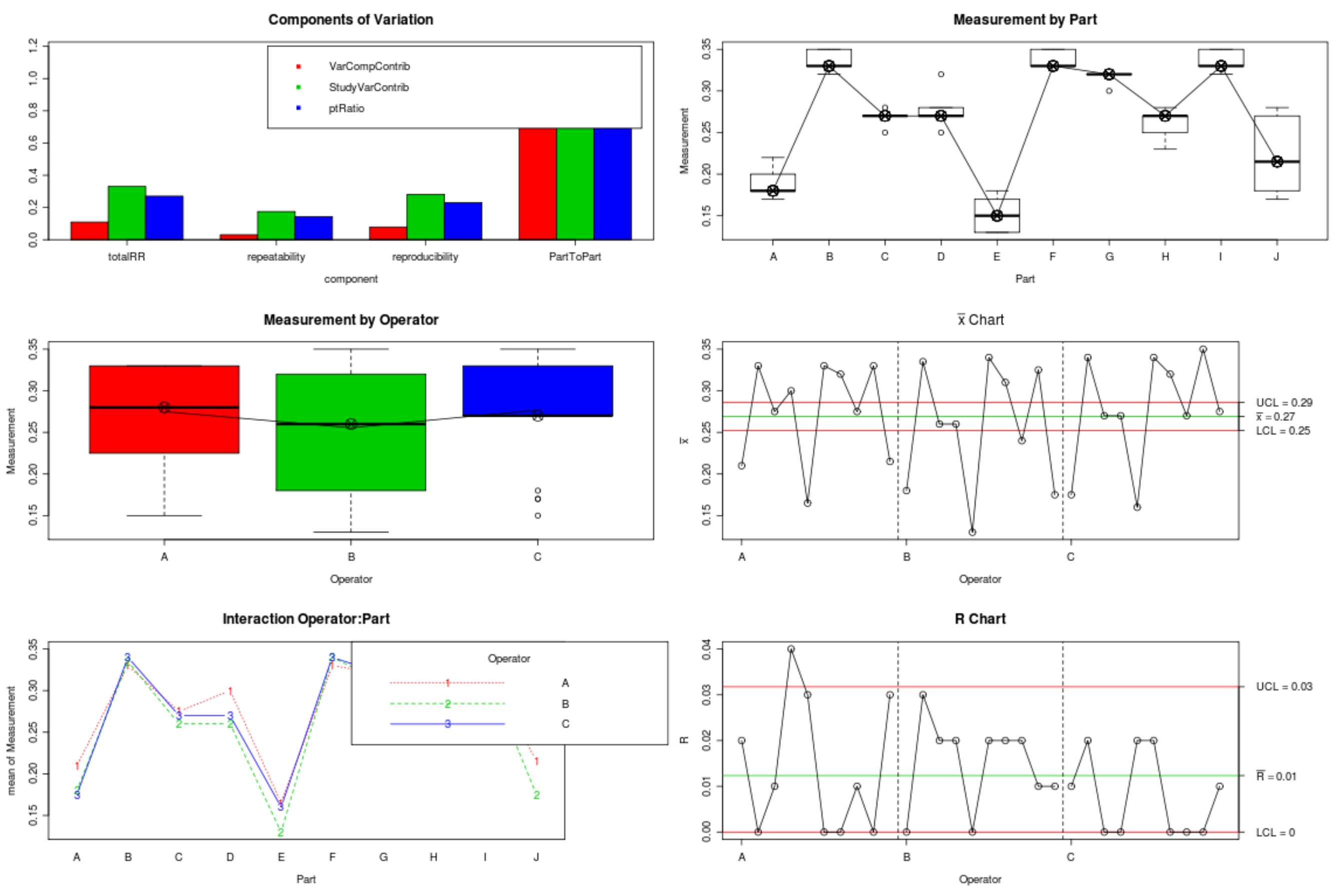
```
Operators:  3  Parts:  10
Measurements:  2  Total:  60
----------

AnOVa Table -  crossed Design
              Df  Sum Sq  Mean Sq F value   Pr(>F)
Operator       2 0.00564 0.002822  19.918 3.13e-06 ***
Part           9 0.22914 0.025460 179.719  < 2e-16 ***
Operator:Part 18 0.01182 0.000657   4.637  0.00011 ***
Residuals     30 0.00425 0.000142
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

----------

Gage R&R
                  VarComp VarCompContrib  Stdev StudyVar StudyVarContrib P/T Ratio
totalRR          0.000507         0.1093 0.0225   0.1352           0.331     0.270
 repeatability   0.000142         0.0305 0.0119   0.0714           0.175     0.143
 reproducibility 0.000366         0.0788 0.0191   0.1148           0.281     0.230
   Operator      0.000108         0.0233 0.0104   0.0624           0.153     0.125
   Operator:Part 0.000258         0.0555 0.0160   0.0963           0.236     0.193
Part to Part     0.004134         0.8907 0.0643   0.3858           0.944     0.772
totalVar         0.004641         1.0000 0.0681   0.4088           1.000     0.818

---
 * Contrib equals Contribution in %
 **Number of Distinct Categories (truncated signal-to-noise-ratio) = 4
```

Notes:

- ⋆ In this example, the Operator by Parts interaction is statistically significant. So, the interaction contributes to Gage variation
- ⋆ The `P/T Ratio` for Gage (`totalRR`) is much larger than 10%, and it is nearly greater than 30%.
- ⋆ On the other hand, the `StudyVarContrib`, for Gage is greater than 30%. This may mean that the measurement system is not acceptable
- ⋆ The main cause for the big Gage variation seems to be attributable to reproducibility, because of large values of Operator and Operator by Part `StudyVarContrib`.
- ⋆ The overall measurerment system should be improved.