PROJECT REPORT

Group 12: Automation in Germany

- ABHAS BHARGAVA - bmat1801
- ARVIND RAGHAVENDRAN - bmat1909
- VARUN PRAMOD AGRAWAL - bmat1940
- ROHAN PRAMOD SHINDE - bmat1925

Topics preceded by a * were not covered in the original project and were done out of self-interest.

PROJECT REPORT

Introduction:

What is automation? Automation is the creation and application of technologies to produce and deliver goods and services with minimal human intervention. The implementation of automation technologies, techniques and processes improve the efficiency, reliability, and/or speed of many tasks that were previously performed by humans.
Common examples include: A spray painting machine, drilling machine, electric screwdriver, electric toothbrush, etc.

In the above examples, we see that it is a simple elevation of a task that required mechanical work to be done by the person wielding the tool, into making a machine do the work for them.

Automation can also be applied to an already automated tool.
For example- An automated storage unit, a robotic spray paint machine used in the industry, etc.

Necessity for Automation:

Why do we want to mechanize our workstation? Global competition and customer demand for high product variety lead to a higher degree of individualization and increase competition among manufacturers which favour new developments like mass customisation. This, in turn, leads to a growing number of product variants. Globalisation also generates enormous cost pressure especially in high-wage countries. High labour costs, increasing production numbers and quality demands require a higher Level of Automation (LoA).

Level of Automation (LoA):

What is Level of Automation? From 2004-2009, The DYNAMO (Dynamic Levels of Automation) project [1] was built, which systematically evaluated LoA by not only studying the machinery, but also the level of interaction between the human and two types of technology mechanization and computerization. Based on that, the LoA is defined as the quotient of the set of functions already automated for the respective process and the set of all necessary functions and scaled according to the DYNAMO project. Quite intuitively, it is the rank of the automated system in the hierarchy of machines in an industry.

Example of LoAs in an assembly process:

Given below is an example of a 7-stage automated environment.

LEVEL	MECHANICAL LoA	EXAMPLE
1	Manual	Physical Strength
2	Static Hand Tool	Screwdriver
3	Flexible Hand Tool	Adjustable Spanner
4	Automatic Hand Tool	Hand drills
5	Static Workstation	Turning Lathe
6	Flexible Workstation	CNC-Machines
7	Fully Automatic	AGV

- Levels 1 to 4 correspond to manual assembly.
- Levels 5 and 6 of the scale can be equated with hybrid assembly, in which the assembly products are manufactured with a combination of manual workstations and automatic machines.
- Level 7 corresponds to a fully automated assembly.

<u>LoAs in a typical industry:</u>

In the previous table, we saw the different mechanical levels of automation that can exist in an industry. However, for the machines to work, there must also be a systematic flow of commands that dictate it to perform its job. In other words, a management system must be in place.

This represents a typical hierarchy in an industry showing the levels of automation:

Field Level → Control Level → Supervisory Level → Planning Level → Management Level

- The field level consists of the parts that do the work, like the tools in the previous table, an actuator, a sensor, etc.
- The control level employs PLCs and PIDs to control the working of these tools, for e.g., heaters at industrial plants have PIDs to that will tell the time at which the heater must be turned off/on as per plant requirements.
- The supervisory level usually consists of Graphical User Interfaces (GUIs) that can be used to supervise the working of a unit, for e.g., remotely turning on/off a water pump at a plant.
- The planning level has the required software to control a single plant, from the raw material to the delivery system of finished products.
- The management level uses the company's integrated management system to control their operations. It is essentially a bunch of computer applications that takes the data sent in from the lower levels and helps in managing the business.

<u>Pros and Cons of Automation:</u>

Automation is a very essential upgrade that one must provide to their business if they wish to scale it higher. The reason? Humans make errors, but a well-dictated machine does not. We can get more accuracy in the process, make multiple varieties of products of superior quality, and the software ensures that the flow of information is systematic, thereby helping management. However, this is not devoid of any flaws. There is always a risk on increasing the unemployment of the State. Another point is that automation is an expensive process. Although this may not technically be a flaw, but if done without proper pre-planning, the returns may not be good. The complexity of automation is what makes choosing the Level of Automation particularly important.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Our Research methodology:

We will introduce the research methodology given to us. A structured and pre-tested questionnaire was developed to capture responses amongst industry representatives. The questionnaire itself is divided into three main sections –Section A "motivational hypotheses", section B "influencing factors "and section C "demographic questions". The complete questionnaire can be found here.

- Section A collects the appraisal on the expected development of automation within the next five years, the perceived importance of the right automation level in assembly systems and the comparison of the importance and practical consideration of monetary and non-monetary factors.

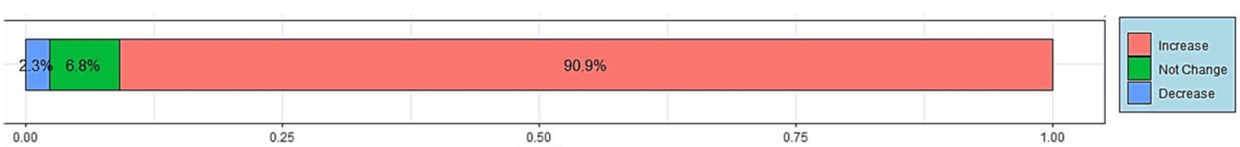
- Section B forms the main part of the questionnaire and captures the participants' opinion on the influence of monetary and non-monetary factors on automation decisions. The list of 19 monetary and 52 non-monetary influencing factors was identified through an extensive literature review. On a first level, they are divided into the three perspectives: "market", "technology" and "monetary" and on a second level into areas.
    - The market perspective contains the areas "market & competitors", "own company", "employees" and "customers".
    - The technology perspective comprises the areas "technology development & production process", "product" and "design".
    - The monetary perspective lists the monetary factors and is not further broken down.

  The data were captured by using a six-point Likert-scale with labeled extremes ranging from "no influence" to "very strong influence". The aim was to prohibit the possibility to take a neutral position but to indicate a tendency in one direction or deliberately refuse the statement.

- Section C captures demographic questions about the characteristics of the companies such as branch or annual turnover and questions about the assembly system of this company such as the current automation level of the assembly system, product type and structure as well as the assembly quantity and personal. Finally, age and gender of the participant are asked for. To assess the current automation level, the seven-point reference scale of the mechanical Level of Automation (LoA) according to the DYNAMO research project is used. Therefore, the scale is shown to the survey participance next to the question.

The survey was published as an online questionnaire via the website www.soscisurvey.de.

A general question was asked to the participants, asking them to give their opinion on how the degree of automation would change in the next 5 years.



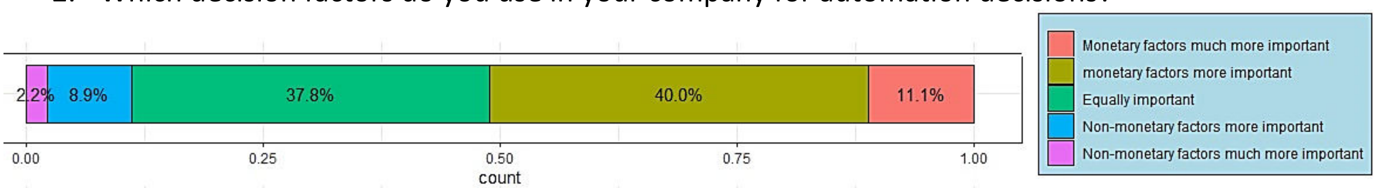
Topics preceded by a * were not covered in the original project and were done out of self-interest.

91% of participants felt that it would definitely increase in the future, which was not very surprising.

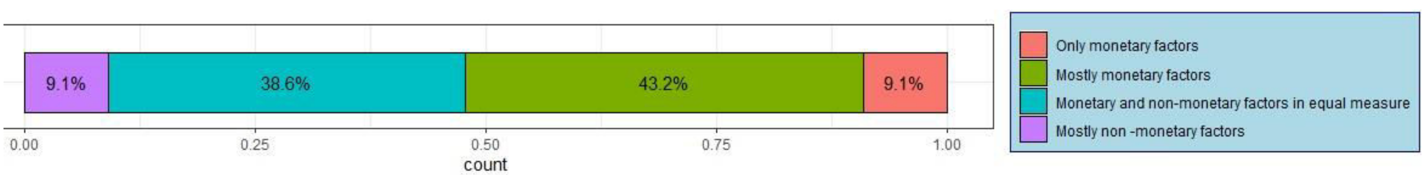
Monetary or Non-monetary?

Traditionally, the ideal degree of automation was determined based on monetary comparative calculations. Labor costs are therefore opposed to automation costs. The expenses for automation are compared with the monetary saving potential. However, research agrees that a purely monetary assessment of the automation level is not adequate.
Two questions related to the same topic was then asked to the participants:

1. Assess the importance of monetary and non-monetary decision factors on the level of automation.
2. Which decision factors do you use in your company for automation decisions?



The results regarding the importance of monetary compared to non-monetary factors in the automation decision indicate a higher importance of monetary factors (51%). Yet, 38% of the representatives, which state that both types of factors are of equal importance, claim that non-monetary factors should gain relevance.



Accordingly, monetary factors are used primarily or solely by 52% of the participants to determine the suitable LoA and 39% of the respondents use both types of factors to the same extent.

These tendencies towards monetary factors can be explained by a lack of methods, which consider non-monetary factors to determine the optimum level of automation of flow-line assembly systems. To close this gap and to facilitate the use of non-monetary factors, the research presented in the paper utilizes those factors to determine a guide value for the LoA, which serves as a basis for further, detailed assessments.

Describing the Flow Assembly System:

We will now proceed to describe a flow assembly system. Consider a typical flow assembly system, then we can divide its properties into four categories:

Topics preceded by a * were not covered in the original project and were done out of self-interest.

- Technical properties of the assembly system, e.g., takt time and buffer time.
- Product properties, e.g., shape, weight, etc.
- Workplace design, e.g., work design measures, ergonomics, muscular and physical strain.
- Logistical influences. e.g., material provision at the assembly line and warehouse.

Each of these description factors were assigned 2-5 characteristics that were determined based on above literature and were also given increments.
For example,

- Spatial Arrangement: U-shape, O-shape, Line
- Annual Production Volume: 1000-5000, 50000-100000, >100000

These characteristics were determined based on the above literature and suitable ranges or increments were identified. The number of characteristics ranges from two to five per description factor. These increments establish a connection between the different factors and a possible/ meaningful automation. Thus, each characteristic is linked with a numerical value that represents the automation suitability. An increasing degree of automation is evaluated with a higher value for this factor. To describe the system, the appropriate characteristic of each description factor is selected. The overall automation rating of the assembly system f_total can then be calculated using the following equation:

$$f_{total} = \frac{\sum_i \left(\frac{\sum_s f_{is}}{S} \right) + \sum_j f_j}{i + j}$$

Since a flow-line assembly consists of several assembly stations, the morphology of the workplace design and some factors (*i*) of the technical and product properties should be evaluated separately for each station *s* and then averaged (first summand in numerator). All factors which are evaluated for the whole assembly (*j*) are also summed up (second summand in numerator). The sum of all automation values is then divided by the total number of factors (*i+j*).

<u>Describing the factors that affect LoA:</u>
Based on the assembly description model and the concept of the overall automation rating f_total, a determination method for a guideline value, called target Level of Automation (LoA_target) was developed, to which the current situation of each assembly process step can be compared to assess potential automation measures. For this purpose, the description factors were extended to also incorporate the production environment. This factor set is named “influencing factors” in the following. They can be subdivided into “market perspective” and “technology perspective”. The “market perspective”, is further divided into the areas of influence “market & competitors”, “own company”, “personnel” and “customers”. The “technology perspective”, comprises “technology & assembly process”, “product” and “product design”.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

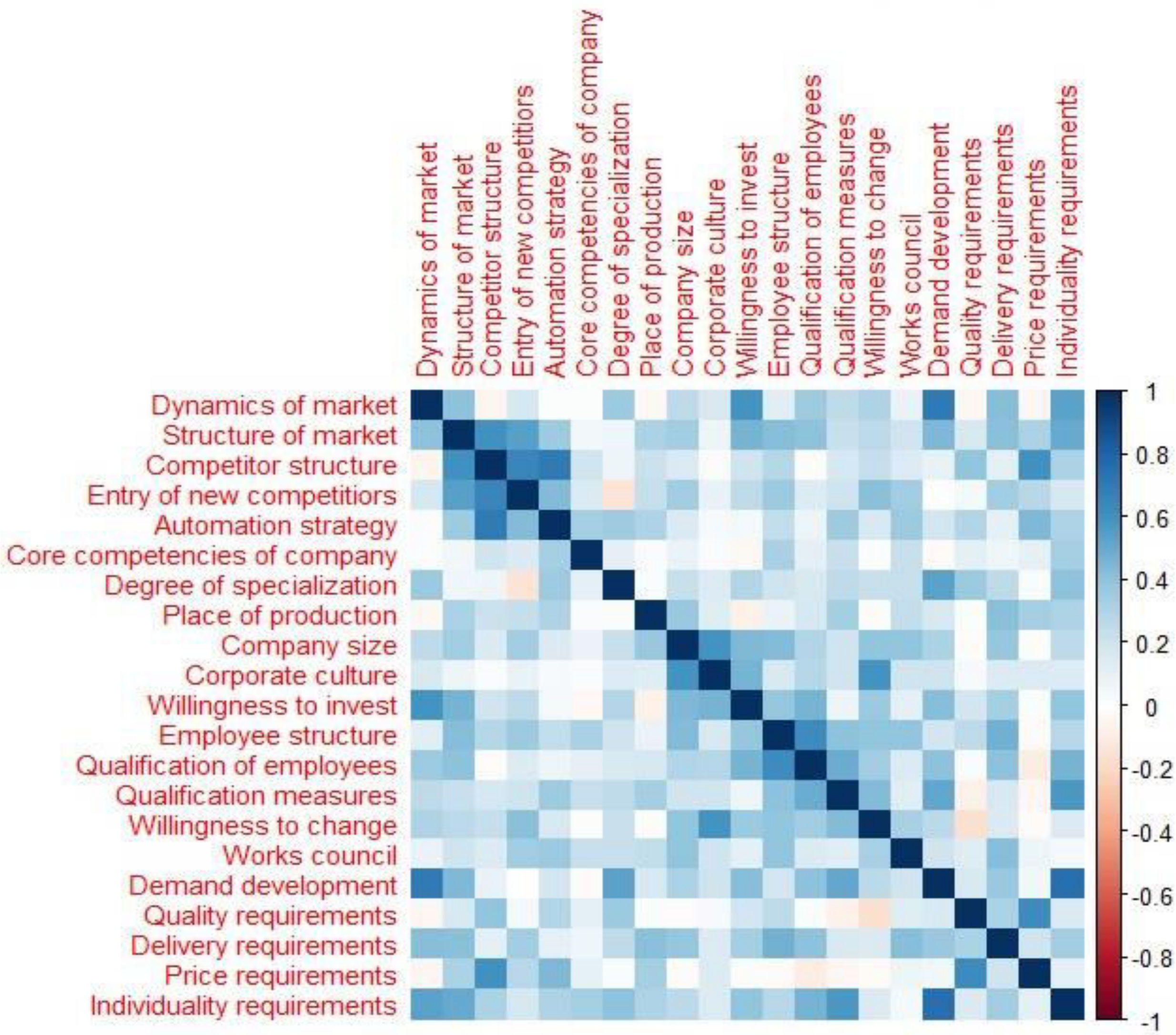
***Market Perspective:**

When we speak about the industry, we cannot avoid talking about the market. So, it makes sense that we look at automation from a market perspective, that is, how do the factors related to the market affect the degree of automation? We can divide this into 4 categories,

- Market and competitors
- Own company
- Personnel
- Customers

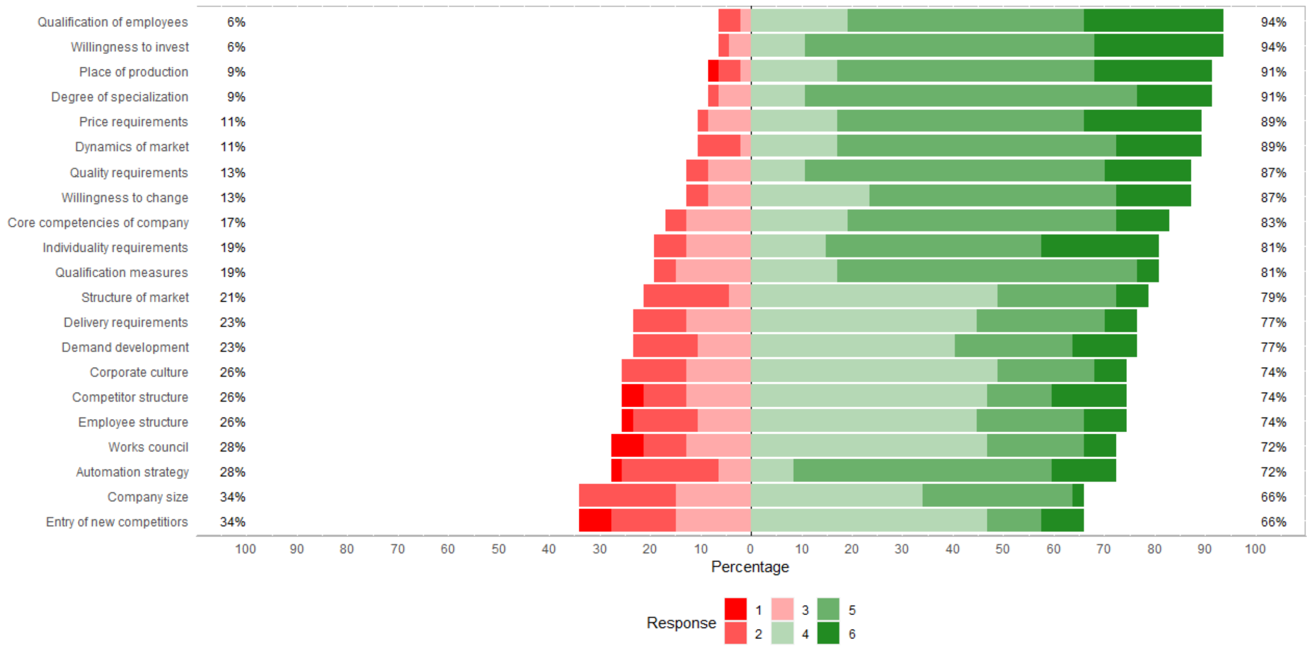
The factors in each of these categories were taken and then the participant was asked to give their opinion on how each of those factors would affect the degree of automation. To improve ergonomics for respondents (i.e., avoiding tiresome scrolling through the questionnaire), the factors were presented in sections in the questionnaire.

The answers were recorded. The correlation plot will help us deduce if a linear relation can exist between any of the factors



Topics preceded by a * were not covered in the original project and were done out of self-interest.

The squares of a darker shade are more inter-related than those of a lighter shade. Let us analyze this by looking at the divergent row chart between the factors.



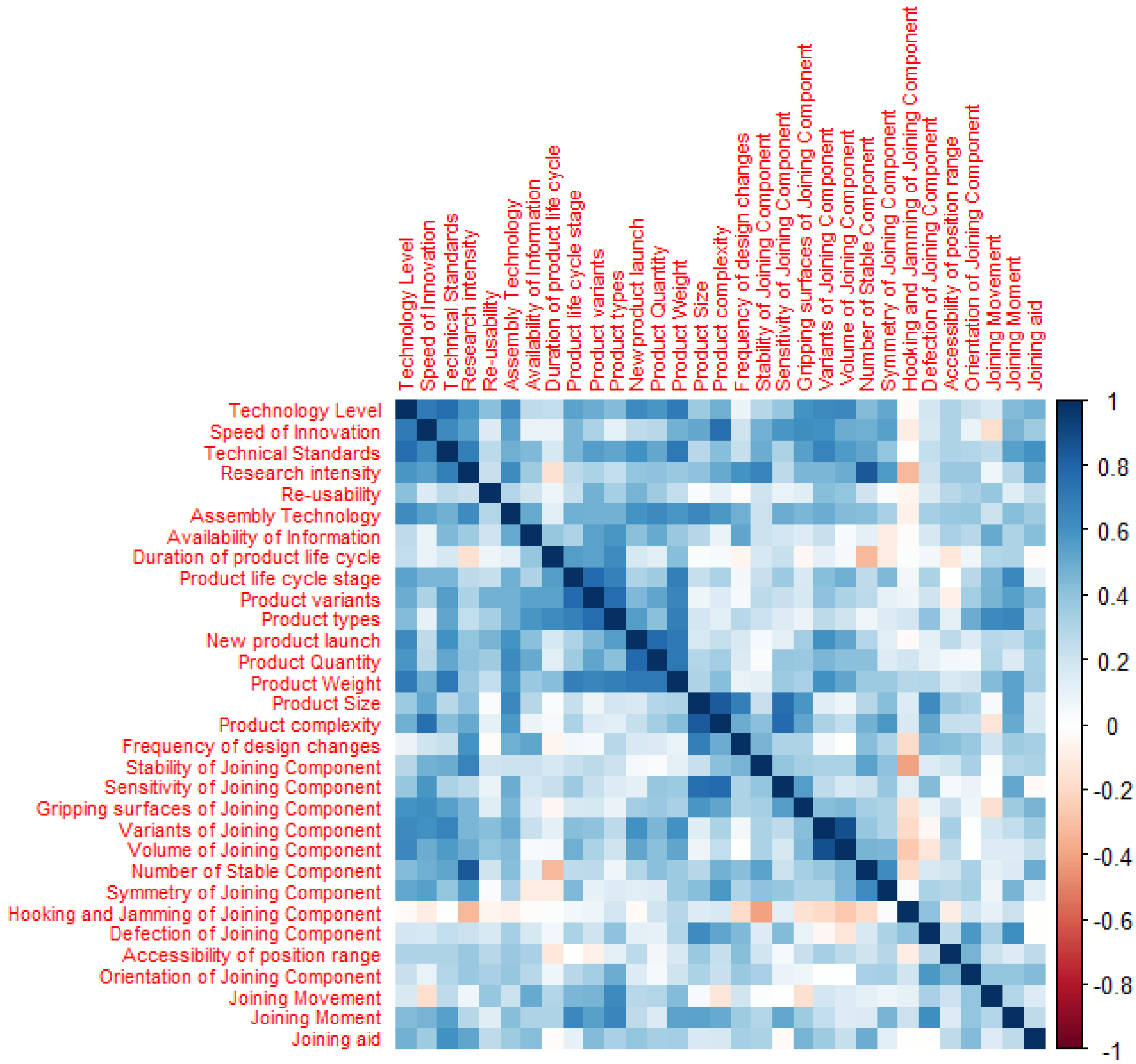
The graph tells us that the participants have a similar opinion towards both the categories. Participants feel that the two categories will either affect the Degree of Automation heavily, or not at all, with a fraction of the participants choosing to remain neutral. The correlation plot is important because it tells us about the validity of the data we obtained. If we want to produce goods of high quality at sustainable rates, then that will naturally lead to a higher degree of automation; so, it makes sense that they are positively correlated, which is also shown in the correlation plot. While this makes sense intuitively, the correlation of a lot of pairs might not make sense just by a glance. However, a brief analysis of previous years' data [2] tells us that the data we have is a good estimate of the scenario in the market.

<u>*Technology Perspective:</u>

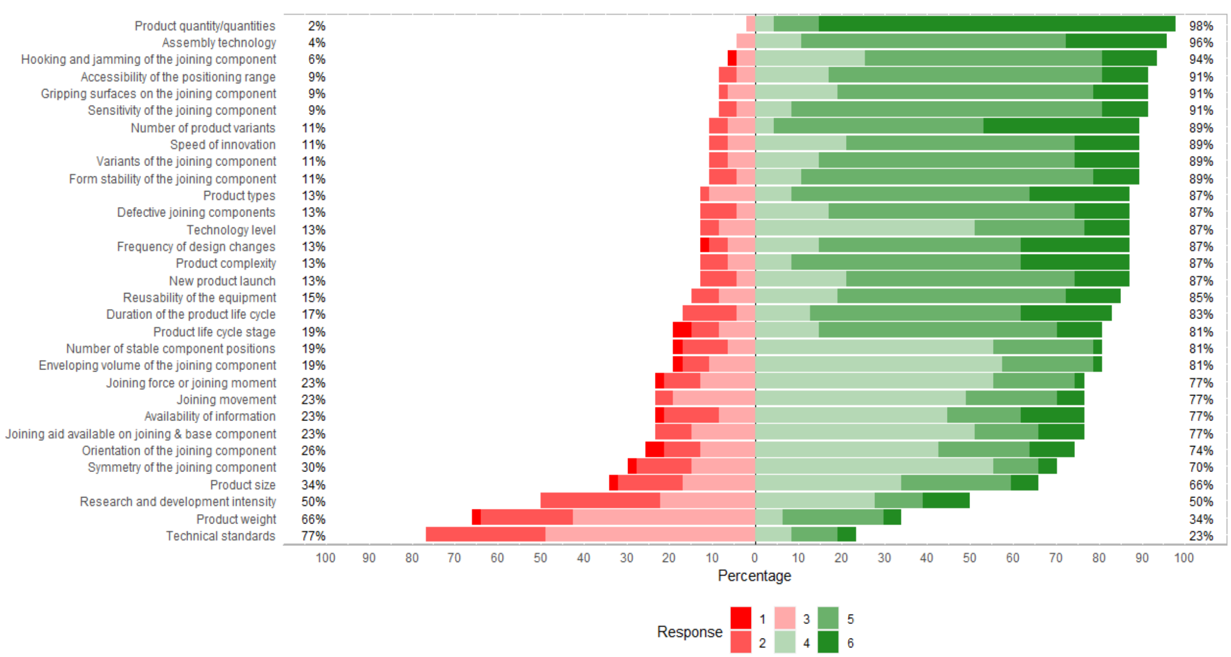
Another factor of great concern is the technological prowess. We can group these factors into three categories,

- Technology development & production process
- Product
- Construction

A similar process as in the case of Market Perspective was done, and we once again look at the correlation plot.



Let us check the validity of the data using the divergent row plot between the factors.



Topics preceded by a * were not covered in the original project and were done out of self-interest.

If customer demand pushes us to increasing the number of types of products, then we will have a new product launch soon. So, it makes sense that "Number of product variants" and "New product launch" are positively correlated, which is also shown from the correlation plot. As mentioned earlier, not a lot of pairs might be intuitively true, but it is a good fit with the market scenario [3].

Computing $\text{LoA}_{\text{target}}$:

Coming back to our research, we now have seven areas of influences to work with: Three from TP and four from MP. Recall that we had seen how we can divide the automation classes into manual, hybrid and automated. For each decision factor, three characteristics correlate with the above-mentioned automation classes "manual", "hybrid" and "automated" assembly with a value f from 1 to 3 being assigned to each characteristic. To determine the LoA_target for an assembly system, the existing value for each factor i must be selected. The class value K is calculated using the following equation:

$$K = \frac{\sum_{i=1}^{52} g_i f_i}{7}$$

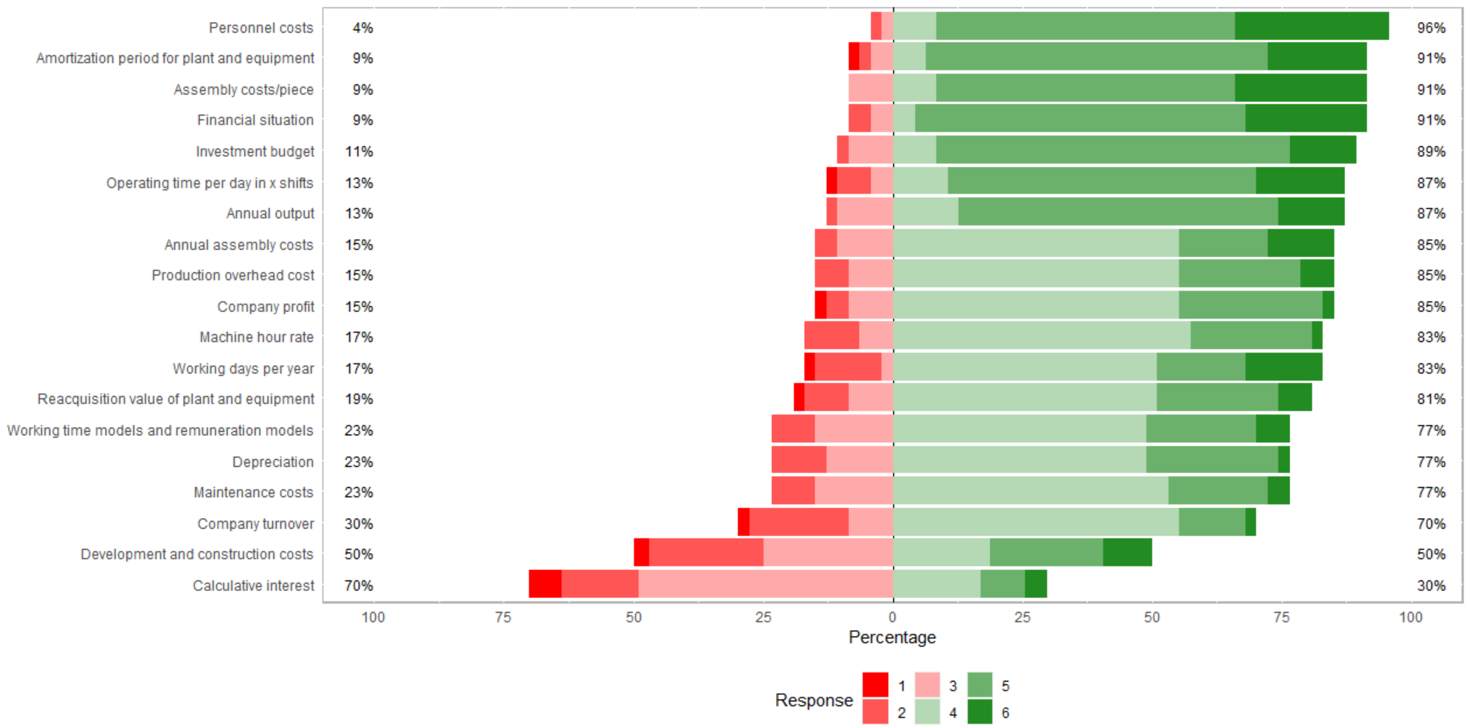
The above formula is applied after taking a weighting g_i of the 52 factors within the seven areas of influences into account, which results as well from the survey and expert interviews. The calculated class value is again between 1 and 3 and is converted into the LoA_target using the given scale:

$\text{LoA}_{\text{target}}$	1	2	3	4	5	6	7
K	1.0-1.1	1.1-1.2	1.2-1.3	1.3-1.4	1.4-1.9	1.9-2.4	2.4-3

***Analysis of Monetary Factors:**

Before moving to the actual cost benefit analysis, it makes sense to talk about the monetary factors. There are no categories into which these factors can be divides, they are just listed, for example, "Personnel Costs", "Financial situation", "Depreciation', etc. The divergent row chart will help us analyze the factors.

As seen in the graph, most of the factors play a significant role in deciding the level of automation in the company. This strongly resonates with the fact that more than half the participants, 53% to be accurate, solely base their automation-related-decisions on monetary factors. "Calculative Interest" and "Development and Construction Costs" play a minor role, with just 30-50% influence.



***Analysis of Likert Data Scales:**

Likert items are used to measure respondents' attitudes to a particular question or statement. Possible items can be “no influence”, “moderate influence”, etc. A Likert scale is composed of a series of four or more Likert-type items that represent similar questions combined into a single composite score/variable. The scale in our project is:

1 No Influence
2 Less Influence
3 Fairly Influential
4 Moderately Influential
5 Strong Influence
6 Very Strong Influence

For our project, since we had a Likert type data, we could not use the raw data itself for the linear regression tests. The data had various non-monetary factors divided into Technology and Market perspective. Each of these categories was then subdivided into 4-5 subcategories. The missing values in the non-monetary factors data was replaced with the median of the respective column because the number of observations at hand was quite small. Replacing the null values in a column with mean was not a promising idea especially since the data is based on 6-point Likert scale. If a portion of the sample prefer extreme values, then the mean can become centralized, which might not be an accurate representation of the population. The most appropriate measure is the median.

A frequency distribution of responses is more helpful. Parametric tests like t tests, regression, etc. can be done as well. In our case, we calculate mean of columns by a single person and then assign it to a new variable and in doing so, we combine 4 or more Likert type items into a single composite score using the mean. However, to describe the data, means are often of limited value unless the data follow a classic normal distribution.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Generally recommended tests [4] are:

- Mann-Whitney U test
- Kruskal-Wallis Test
- Chi Square Test

We will elaborate in detail about the Linear test and Chi Square test.

*Linear Regression tests

After importing the dataset in R and then removing unwanted columns we chose a particular category to work with like Market Perspective and the technology Perspective. A new data frame is then created with columns as the subcategories of each category above e.g., Market category was subdivided into Market competitors, Own company, Personnel and Customers while the Technology perspective was divided into Construction, Technology and Product. The values in this new data frame were obtained by taking the mean of values in respective row of the columns in the columns that fall under the respective subcategory. Now since the values in this data frame are continuous (since they are mean) we can apply the linear regression on this data frame with the LoA column appended to this data frame to study and discuss the linear relationship between given LoA and the factors in the new data frame if any.

The first rows of this data frame for Market perspective would look something like this:

	LoA	Market.competitors	Own.company	Personnel	Customers
1	5	2.4	3.5	3.2	3.8
2	6	3.4	4.0	3.2	4.4
3	6	4.4	4.0	3.4	4.8
4	7	2.0	4.2	4.0	3.8
5	6	6.0	5.5	6.0	6.0
6	3	2.8	4.5	3.8	3.8
7	6	5.2	3.5	4.0	4.6
8	4	5.0	5.0	5.0	5.0
9	6	3.6	4.2	4.4	4.6
10	3	4.4	4.7	5.0	5.0
11	6	4.4	4.8	3.6	4.4
12	5	4.6	5.0	4.8	5.4
13	5	4.0	4.3	3.0	5.0
14	7	4.0	3.5	4.6	4.8
15	3	4.4	4.7	4.4	4.2
16	5	3.4	4.2	3.6	3.6

Similar type of data frame can be made for Technology perspective.

The next step we deal with is forming a linear model on both data frames to study a linear relationship between LoA and other factors. The linear model is formed as follows:

```
> mp.lm <- lm(LoA~., mp)
> mp.final.lm <- step(mp.lm)
```

The step function performs an ANOVA test on the independent factors and then checks which of the factors have a high chance to be linearly related to LoA column. It then makes a new linear model based on the relevant independent factors obtained after repeatedly performing the cycle until no more relevant factors can be obtained from the model based on ANOVA tests. Note that here, the irrelevant factors are based on a 5% confidence interval.

Calling the summary function on the model provides the statistical information on the linear model like the p-value with the null hypothesis being that the factors have a coefficient 0 in the linear model, p-value for each factor, F-statistic, etc.

The summary performed on subcategories of Market perspective is as follows:

```
> mp2.lm <- lm(LoA~., mp2)
> summary(mp2.lm)

Call:
lm(formula = LoA ~ ., data = mp2)

Residuals:
    Min      1Q  Median      3Q     Max
-3.1165 -0.4271  0.2054  0.8597  1.6804

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)          7.1314     1.8339   3.889 0.000566 ***
Market.competitors  -0.3732     0.3198  -1.167 0.253074
Own.company         -1.2020     0.5378  -2.235 0.033562 *
Personnel            0.2715     0.3649   0.744 0.463109
Customers            0.8342     0.3937   2.119 0.043120 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.303 on 28 degrees of freedom
Multiple R-squared:  0.2363,	Adjusted R-squared:  0.1271
F-statistic: 2.165 on 4 and 28 DF,  p-value: 0.09898
```

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Each of the values in the summary have a particular meaning as stated in the class.

We provide a table to find if the linear model is the best fit for the data:

STATISTIC	CRITERION
R-Squared	Higher the better *(> 0.70)*
Adj R-Squared	Higher the better
F-Statistic	Higher the better
Std. Error	Closer to zero the better
t-statistic	Should be greater 1.96 for p-value to be less than 0.05
AIC	Lower the better
BIC	Lower the better
Mallows cp	Should be close to the number of predictors in model
MAPE (Mean absolute percentage error)	Lower the better
MSE (Mean squared error)	Lower the better
Min_Max Accuracy => mean(min(actual, predicted)/max(actual, predicted))	Higher the better

Based on the summary of the model, the null hypothesis can be rejected for the Own company and Customers factors with 95% confidence. For better model performing a slight change in the code as follows we intend to get a better model fit using the step function as described earlier:

```
> mp2.lm <- lm(LoA~., mp2)
> mp2.final.lm <- step(mp2.lm)
```

The summary can be given for this model as:

```
> summary(mp2.final.lm)

Call:
lm(formula = LoA ~ Own.company + Customers, data = mp2)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.2037 -0.6819  0.1875  1.0277  2.0456 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   7.4668     1.8009   4.146 0.000255 ***
Own.company  -1.2402     0.4818  -2.574 0.015231 *  
Customers     0.7095     0.3475   2.042 0.050052 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.295 on 30 degrees of freedom
Multiple R-squared:  0.1921,	Adjusted R-squared:  0.1382 
F-statistic: 3.566 on 2 and 30 DF,  p-value: 0.0408
```

Based on the given summary by R and the previous table, it is quite clear that the linear model is not a correct model fit for the given data. An intuitive explanation is that we had replaced the NA values by median and then taken their mean for the columns of the subcategory, but that might have introduced a lot of asymmetry in the model to fit accordingly. Also, since the data provided is Likert type, it is usually advised to not to try the linear regression on Likert data.

To check if the first point could be the reason for bad fit, we perform the linear model test on original columns of Market perspective itself. It is done as follows:

Topics preceded by a * were not covered in the original project and were done out of self-interest.

```
> summary(mp.final.lm)

Call:
lm(formula = LoA ~ MS02_01 + MS02_02 + MS02_11 + MS03_12 + MS03_14 +
    MS03_16 + MS03_17 + MS04_20 + MS04_19 + MS04_18 + MS04_22 +
    MS05_03 + MS05_05, data = mp)

Residuals:
    Min      1Q  Median      3Q     Max
-1.1000 -0.5033 -0.1648  0.5766  1.2657

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.4013     1.4856   2.289  0.03367 *
MS02_01       0.2913     0.2653   1.098  0.28581
MS02_02       0.4119     0.1832   2.249  0.03658 *
MS02_11      -0.4212     0.1468  -2.869  0.00982 **
MS03_12       0.4292     0.1674   2.565  0.01896 *
MS03_14      -0.2736     0.1533  -1.785  0.09019 .
MS03_16      -0.5816     0.1635  -3.557  0.00210 **
MS03_17      -0.4053     0.2606  -1.556  0.13631
MS04_20      -0.6638     0.2000  -3.320  0.00360 **
MS04_19       0.6105     0.2486   2.456  0.02386 *
MS04_18       0.4339     0.2468   1.758  0.09485 .
MS04_22       0.1987     0.1376   1.444  0.16494
MS05_03      -0.3468     0.2472  -1.403  0.17686
MS05_05       0.5992     0.1749   3.425  0.00284 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9172 on 19 degrees of freedom
Multiple R-squared:  0.7432,    Adjusted R-squared:  0.5675
F-statistic:  4.23 on 13 and 19 DF,  p-value: 0.002328
```

Based on the summary some of the factors seem to be quite linearly related to LoA like **MS02_11** (Automation strategy of competitors), **MS03_16** (Corporate culture), **MS04_20** (Employee structure) and **MS05_05** (Quality requirements).

But as stated earlier the data is ordinal and performing Linear regression on the data is not particularly useful. Similar summaries were obtained for the Technology perspective:

Topics preceded by a * were not covered in the original project and were done out of self-interest.

```
> summary(tp2.lm)

Call:
lm(formula = LoA ~ ., data = tp2)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.2598 -0.9869  0.3844  0.8539  2.2115 

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)   
(Intercept)              5.1016     1.7074   2.988  0.00567 **
Construction             0.2846     0.4486   0.635  0.53068   
Product                 -0.3691     0.4117  -0.897  0.37735   
Technology.development   0.1240     0.4452   0.278  0.78262   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.443 on 29 degrees of freedom
Multiple R-squared:  0.03032,	Adjusted R-squared:  -0.06999 
F-statistic: 0.3023 on 3 and 29 DF,  p-value: 0.8235
```

```
> summary(tp.final.lm)

Call:
lm(formula = LoA ~ TS04_20 + TS04_21 + TS04_22 + TS04_24 + TS04_26 + 
    TS04_27 + TS04_29 + TS04_30 + TS04_31 + TS04_32 + TS03_05 + 
    TS03_06 + TS03_07 + TS03_08 + TS03_09 + TS03_10 + TS03_11 + 
    TS03_12 + TS03_13 + TS03_14 + TS02_03 + TS02_04 + TS02_15 + 
    TS02_16 + TS02_18, data = tp)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.51792 -0.11276  0.01558  0.15525  0.48308 
```

Topics preceded by a * were not covered in the original project and were done out of self-interest.

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.7675     1.3724   5.660 0.000767 ***
TS04_20        2.3744     0.4463   5.320 0.001100 **
TS04_21       -4.6587     1.0925  -4.264 0.003727 **
TS04_22        6.4201     1.3262   4.841 0.001876 **
TS04_24       -4.1925     0.7247  -5.785 0.000674 ***
TS04_26       -1.6274     0.5017  -3.244 0.014174 *
TS04_27        0.2660     0.2239   1.188 0.273471
TS04_29       -1.0261     0.2542  -4.037 0.004954 **
TS04_30        0.3399     0.2407   1.412 0.200808
TS04_31        0.5820     0.2121   2.744 0.028766 *
TS04_32        3.8225     0.8536   4.478 0.002873 **
TS03_05       -2.9406     0.5978  -4.919 0.001715 **
TS03_06       -0.9247     0.3428  -2.698 0.030748 *
TS03_07        3.9177     0.6657   5.885 0.000608 ***
TS03_08       -2.6606     0.3765  -7.066 0.000200 ***
TS03_09        3.5482     0.5950   5.964 0.000562 ***
TS03_10       -0.7006     0.3078  -2.277 0.056929 .
TS03_11       -0.5559     0.2550  -2.180 0.065648 .
TS03_12        1.5732     0.3063   5.136 0.001345 **
TS03_13       -3.9893     0.7376  -5.409 0.000999 ***
TS03_14       -2.2764     0.4094  -5.561 0.000850 ***
TS02_03       -0.6147     0.1539  -3.993 0.005235 **
TS02_04        2.3082     0.4818   4.790 0.001988 **
TS02_15       -0.6812     0.2419  -2.816 0.025939 *
TS02_16        1.1476     0.2791   4.112 0.004507 **
TS02_18        0.3182     0.1452   2.191 0.064602 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4853 on 7 degrees of freedom
Multiple R-squared:  0.9735,    Adjusted R-squared:  0.8789
```

```
F-statistic: 10.29 on 25 and 7 DF,  p-value: 0.001959
```

So, even the factors from the Technology Perspective are not linearly related to the LoA.

<u>*Chi Squared Association tests</u>

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Independence test between LoA and Assembly quantity

```
> df$DF02_01 <- as.factor(df$DF02_01)
> df$DF03_01 <- as.factor(df$DF03_01)
> tbl1 <- table(df$DF02_01,df$DF03_01)
> chisq.test(tbl1)

        Pearson's Chi-squared test

data:  tbl1
X-squared = 48.612, df = 35, p-value = 0.06284

Warning message:
In chisq.test(tbl1) : Chi-squared approximation may be incorrect
>
```

From the chi squared test, we see that the p-value is greater than 0.05 thus we cannot reject the null hypothesis that they are independent with 95% confidence. Thus, the level of automation and Assembly quantities seem independent with 95% confidence.

Independence test between Importance of Non-monetary factors and Usage of Non-monetary factors in a company

```
> df$MH04_01 <- as.factor(df$MH04_01)
> df$MH05_01 <- as.factor(df$MH05_01)
> tbl2 <- table(df$MH04_01,df$MH05_01)
> chisq.test(tbl2)

        Pearson's Chi-squared test

data:  tbl2
X-squared = 49.904, df = 12, p-value = 1.452e-06

Warning message:
In chisq.test(tbl2) : Chi-squared approximation may be incorrect
>
```

The test results give a quite low p-value (about order of 10^{-6}) and hence even at 99% confidence we can reject the null hypothesis, so these motivational hypothesis factors are clearly not independent.

Independence test between Level of Automation and Annual turnover of the company in 2017

```
> df$DF02_01 <- as.factor(df$DF02_01)
> df$DF07_01 <- as.factor(df$DF07_01)
> tbl3 <- table(df$DF02_01,df$DF07_01)
> chisq.test(tbl3)

        Pearson's Chi-squared test

data:  tbl3
X-squared = 30.334, df = 25, p-value = 0.212

Warning message:
In chisq.test(tbl3) : Chi-squared approximation may be incorrect
> 
```

Here the p-value is 0.212 which is much greater than even 0.1. Thus, even at 90% confidence we cannot reject the null hypothesis of independence. Thus, LoA clearly depends on Turnover of the company.

Independence test between LoA and Number of Assembly workers in the company

```
> df$DF02_01 <- as.factor(df$DF02_01)
> df$DF05_01 <- as.factor(df$DF05_01)
> tbl4 <- table(df$DF02_01,df$DF05_01)
> chisq.test(tbl4)

        Pearson's Chi-squared test

data:  tbl4
X-squared = 27.471, df = 30, p-value = 0.5985

Warning message:
In chisq.test(tbl4) : Chi-squared approximation may be incorrect
> 
```

Here too the p-value 0.5985 is quite greater than 0.1 and hence we can reject the null hypothesis of independence. So LoA is clearly dependent on Number of Assembly workers which is quite intuitively correct.

Independence test between LoA and the industry the company belongs to.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

```
> df$DF02_01 <- as.factor(df$DF02_01)
> df$DF06 <- as.factor(df$DF06)
> tbl5 <- table(df$DF02_01,df$DF06)
> chisq.test(tbl5)

        Pearson's Chi-squared test

data:  tbl5
X-squared = 22.977, df = 20, p-value = 0.2899

Warning message:
In chisq.test(tbl5) : Chi-squared approximation may be incorrect
> 
```

Here too the p-value 0.2899 is greater than 0.1 and hence even at 90% confidence interval we can reject the null hypothesis of independence. Thus, the Level of automation and the Industry to which the company belongs to is quite dependent on each other. This can be intuitively justified because a certain type of industry may require a higher level of automation while some other industry might prefer more manual work than automated machinery.

Independence test between LoA and Assembly costs per piece (a Monetary factor)

```
> df$DF02_01 <- as.factor(df$DF02_01)
> df$MF01_11 <- as.factor(df$MF01_11)
> tbl6 <- table(df$DF02_01,df$MF01_11)
> chisq.test(tbl6)

        Pearson's Chi-squared test

data:  tbl6
X-squared = 25.649, df = 15, p-value = 0.04188

Warning message:
In chisq.test(tbl6) : Chi-squared approximation may be incorrect
> 
```

The test results give a p-value less than 0.05 and hence at 95% confidence we can reject the null hypothesis, so the Level of automation in the company and the Assembly cost per piece may be assumed to be dependent on each other with 95% confidence.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Cost Benefit Analysis:

Both, effort, and benefit are often expressed as monetary values to be able to easily identify the tipping point where benefits exceed costs. To be able to include non-monetary factors, we chose a dimensionless scale instead. The point of this analysis was to show the importance of non-monetary factors. The cost-benefit analysis for this paper requires designing equipment scenarios for all relevant LoA. The first step is to break down the assembly process into assembly steps, where each assembly step can be of type handling, joining, or testing. In the next step, the planner attributes an LoA according to Table 1, called LoA_as-is to each assembly step. At this stage, the planner has complete knowledge of the assembly system in question including the current automation situation. After this, the task is to compare possible alternatives with the as-is state. To do so, a team of process experts must make rough concept plans for different equipment scenarios, starting with a scenario that matches LoA_target. In the example in table, "tighten four screws", equipment scenarios could be "screwdriver with bit set", "electric screw wrench" or "automated screwing station". The information needed in each concept plan must follow a predefined structure to ensure all input for the benefit and cost assessment is available. Each equipment scenario is then subject to the assessment steps described hereafter. The planning approach is supposed to be iterative and decision quality improves with larger numbers of scenarios planned.

Benefit assessment:

Before talking about the benefit assessment, we talk a little about the process used, called Analytical Hierarchy Process [5]. Benefit is something that cannot be computed quantitatively, since it has factors like flexibility, quality, etc. and these are comparative with respect to each other; so, this process is employed. We define the objective, structure elements in criteria, sub-criteria, alternatives etc., make a pair wise comparison of elements in each group, calculate weighting and consistency ratio, and then evaluate alternatives according weighting. This may sound simple, but the process involves construction of matrices of the weights and calculation of the normalized eigenvectors, and then analyzing each entry with the computed values, to establish a ranking. The implications of the AHP are more important than the process itself.

A benefit assessment is intended to provide managers with a basis for decision making and it also provides funders with answers to the question of what they are getting for their money.
For doing benefit assessment, six relevant target dimensions were identified for the consideration of the benefits:

- Flexibility – Tells us within what time frame we can get the job done.
- Quality – Tells us the quality of the product
- Productivity – Tells us about the production per unit time. The appropriate number of certain machineries also depends upon this
- Availability – If the product is not available to people at time, they would prefer not to buy the product from them in future.
- Costs – If there are higher making costs than making the product manually the latter would be preferred.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

- Health, Safety and Environment – If there are safety issues for the neighboring areas it is a problem.

These dimensions are not quantitative since we cannot establish a rigid link between them and monetary benefits. The AHP is done, i.e., using a pairwise comparison the user puts all target dimensions in relation to each other to calculate a standardized weighting vector.

Now for this factor we can calculate the benefit changes in relation to the current benefit state of the LOA and we call it as B. It can be experimentally seen that $-6 \le B \le 6$ for each target dimension. This bound is derived from the AHP method.

$B > 0$ indicates an improvement over the existing LOA, while $B < 0$ values signify a deterioration compared to the as-is state. By default, $B = 0$ for $LoA_{as\text{-}is}$ since we consider the benefit of changing the current LoA to any other LoA. This process is equally suitable for reducing the LoA as well. Any equipment that is discarded in the process is treated as a benefit.

So, this assessment provides a dimensionless utility variable (B) that indicates the quality of the expected adaptation of an assembly step from $LoA_{\mathbf{as\text{-}is}}$ to any other LoA. Here, quality is the benefit of increasing/decreasing the LoA

Cost assessment:

Why do we need to perform cost analysis if we can do benefit analysis? The reason is that this data is more quantifiable. This data can be expressed as monetary factors.

Cost can be defined as consumptions or efforts that need to be delivered for implementing the automation adaptation measures – i.e., how much it costs to increase the current level of automation.

For the various equipment scenarios, cost analyses are performed. This data can be expressed as monetary factors, i.e., it is quantifiable.

Cost: Consumptions or efforts that need to be delivered for implementing the automation adaptation measures.

Costs are of 4 types:

- Acquisition Costs
- Planning Costs
- Personnel Costs
- Start-Up Costs

These cost types need to be gathered or estimated for the equipment scenario of each relevant LoA. As with the benefit assessment, costs for the $LoA_{\mathbf{as\text{-}is}}$ are defined as zero.

The following computation is done:

$$C_{z,\,i} = C_{A,\,i} + C_{S,\,i} + C_{P,\,i} + C_{R,\,i}$$

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Where, in the above formula,

z	LoA ; $z \in \{1, \ldots, 7\}$
i	Assembly step
$C_{z,i}$	Total cost per LoA *z* at assembly step *i*
$C_{A,i}$	Acquisition cost at assembly step *i*
$C_{S,i}$	Personnel cost (staff) at assembly step *i*
$C_{P,i}$	Planning cost at assembly step *i*
$C_{R,i}$	Start-Up Cost (ramp-up) at assembly step *i*

$C_{z,i}$ can take arbitrary values between 1 to 10^7 (possibly even higher), so direct comparison is harder. We aim to convert the costs into a dimensionless utility value, like in the case on benefit analysis. We scale the costs by computing the normalized cost value for LoA z at assembly step i, $C_{z,i,norm}$:

$$C_{z,i,norm} = C_{z,i} * 6 / \max\{C_{z,i} \mid z \in \{1, \ldots, 7\}\}$$
$$C = C_{z,i,norm}$$

The purpose of scaling cost at all LoA *z* is to obtain costs for equipment scenarios relative to the worst possible scenario, which will take the value 6. We have now found a value like Benefit(B), i.e., Cost(C) such that, $0 \le C \le 6$.

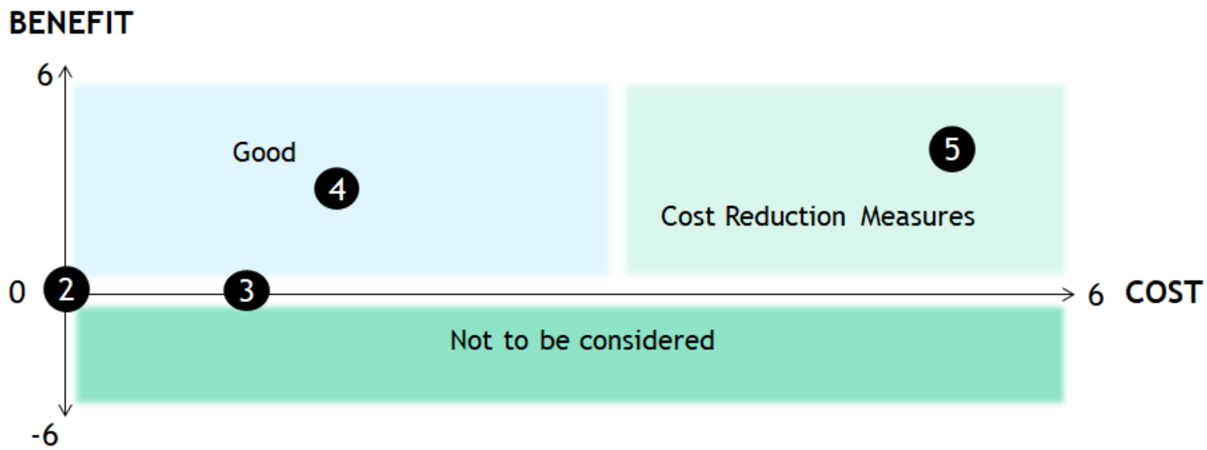
Selecting the optimal LoA:

We plot (C, B) pairs and try to pick an optimal LoA. By default, (C, B) = (0,0) for $LoA_{as\text{-}is}$.
All other planned equipment scenarios for the different LoA will receive value pairs within the ranges $0 \le C \le 6$ and $-6 \le B \le 6$. Planning and evaluating all 6 LoA for each assembly step of the assembly system is tedious.
So, the process is done in the following steps:

- First done for LoA_{target}
- Then for an interval around LoA_{target}, namely, $LoA_{target}-1$, LoA_{target}, $LoA_{target}+1$
- Finally, for all

An example at LoA=2:

We consider the example of screw-driving process, which has $LoA_{as\text{-}is}=2$, by Table 1.
With a determined LoA_{target} of 4, scenarios were planned for LoA 3, 4 and 5 and put through the benefit and cost analyses. The following diagram is considered while explaining this example.



2	Standard Screwdriver	$LoA_{as\text{-}is}=2$
3	Screwdriver with bit set	$LoA_{target}-1=3$
4	Electric screw wrench	$LoA_{target}=4$
5	Automated screwing station	$LoA_{target}+1=5$

Anything below the x-axis is neglected, since it has negative benefit and is not good for the company. From the figure, we see that 4 and 5 have a positive benefit. Now although 5 has a higher benefit, 4 is optimal since it is cost efficient as well. However, with sufficient cost reduction measures, 5 can be implemented. It depends on the industry's long-term goal.

<u>General Discussion:</u>
If the benefit analysis yields multiple situations with equal benefit(B), then the most viable equipment scenario is decided by the associated cost(C). So, the planning process per assembly step should therefore continue until one scenario can be located as far to the top left corner in the figure. By building the reference LoA using company-level factors, suitable measures to tackle dynamic changes in the company's environment should be easy to identify.

There are certain drawbacks to this process:

- The idea to treat both benefit and cost relative to each other and to the current state instead of zero may seem unintuitive at first.
- Since we are focusing on non-monetary factors, no discrete statements about profitability are possible although proxies are inherently present in the benefit assessment. So, it is recommended to perform amortization calculations before final decision-making.
- The high up-front planning work to use this tool is high. For most accurate results, this method requires equipment scenarios for 6 LoA for *n* assembly steps, which would be very tedious to compute for a complex assembly system.

Topics preceded by a * were not covered in the original project and were done out of self-interest.

Conclusion:

This paper describes a methodology that builds upon three partial models:

- An assembly line description to estimate a LoA_{target} for the assembly line
- A cost and benefit analysis using AHP to evaluate equipment scenarios for LoA_{target} in comparison to the as-is state
- A decision-making model to select suitable configurations.

Traditionally, automation decisions mainly consider monetary factors. In contrast to existing practice and research, the presented approach includes non-monetary factors in the decision-making process when choosing automation decisions. The methodology also allows a varying level of detail and therefore adaptable effort in the evaluation process. It could provide recommendations for new adaptations based on scenarios that have already been considered or even successfully implemented elsewhere.

References:

[1] Validation of the DYNAMO methodology for measuring and assessing Levels of Automation.

[2] Kakati, M., 1997. Strategic Evaluation of Advanced Manufacturing Technology. Int J Production Economics, p. 141–156.

[3] Granell, V., Frohm, J., Winroth, M., 2006. Controlling Levels of Automation – A Model for Identifying Manufacturing Parameters. IFAC Proceedings Volumes *39*, p. 65–70.

[4] Analysing and Interpreting Data from Likert-Type Scales (Gail M. Sullivan, Anthony R. Artino Jr).

[5] Using the analytic hierarchy process (AHP) to select and prioritize projects in a portfolio(Vargas, Ricardo Viana).

Supplementary Books for Reference:

1. Discrete Data Analysis with R (by Michael Friendly, David Meyer)
2. Data Mining with R (by Luis Torgo)
3. An Introduction to Statistical Learning (by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani)
4. The Elements of Statistical Learning-Data Mining, Inference and Prediction (by Trevor Hastie, Robert Tibshirani, Jerome Friedman).

Topics preceded by a * were not covered in the original project and were done out of self-interest.

5. Relations Between Parameters/Performers and Levels of Automation (Asa Fasth, Johan Stahre and Jorgen Frohm)

6. Measuring and Analysing levels of Automation in an Assembly system (Asa Fasth, Johan Stahre, Kerstin Dencker)

7. Statistical Methods for analysing discrete and categorical data recorded in performance analysis (Alan Nevill, Mike Hughes, Stephen Mark-Cooper)

Additional websites referred to

1. www.sthda.com
2. www.stats.stackexchange.com
3. www.rpubs.com
4. www.tutorialspoint.com
5. www.towardsdatascience.com
6. www.edureka.com
7. www.statisticsbyjim.com
8. www.datascienceplus.com
9. www.dataindeed.io
10. https://youtu.be/J4T70o8gjlk