Ca' Foscari
University
of Venice

Single Cycle Degree programme

in Economics, Market
and Finance

Final Thesis

Statistical analysis of KPI in consultancy practices. A real case study on selected clients of fischer Consulting Italia.

Supervisor
Ch. Ma Prof. Maria Barbati

Graduand
Giacomo Negri
Matriculation Number 891623

Academic Year
2023 / 2024

INDEX

INDEX OF THE FIGURES

INDEX OF THE TABLES

Introduction

In an era dominated by the proliferation of data, the ability to handle, manipulate, and extract meaningful insights from this vast source of information is crucial for enterprise success. In a more competitive business environment than ever, with a globalized and interconnected world, being able to make impactful decisions in a more effective way is a key to success. However, amidst the sea of numbers and observations, it's easy to get lost, and the sheer volume of data can be overwhelming. Furthermore, data, in its raw form, is merely a recording, not always an accurate reflection of real observations. The true value of data emerges when we can uncover the hidden knowledge within and decipher the relationships it holds. Yet, this is no simple task, as the famous American 19th century writer, Mark Twain, astutely observed: "There are lies, damned lies, and statistics," emphasizing the malleability of data to support various claims. When handled adeptly, however, data can illuminate paths to insights that neither individual humans nor raw data alone could provide.

Thousands of computations necessitate either a multitude of human hours or a single computer. The advent of technological tools, such as R, Orange Data Mining, Python, and Stata, has revolutionized the world and, notably, the field of statistics. Nowadays, these statistical software tools enable the automation and swift execution of analyses, turning complex computations into rapid outcomes. While these analyses still require interpretation to empower businesses and gain a competitive edge, they eliminate reliance on mere "nosology" or the tendency to depend solely on assumptions or experiences. Instead, decision processes should be guided by data, observations, and objective analyses.

Given these considerations, consulting firms like fischer Consulting Italia S.a.s., specializing in enhancing industrial processes and organizational operations, recognize the paramount importance of managing and understanding client data. Their belief lies in leveraging clients' abilities and tools to enhance their performance and meet market and internal needs efficiently.

This thesis delves into the statistical relevance of select consultancy practices on Key Performance Indicators (KPIs) for clients of fischer Consulting Italia S.a.s. (fCIT).

The primary objective of this analysis is to evaluate the potential for fischer Consulting Italia S.a.s. to gain insights into performance differences pre- and post-consultancy and, potentially, to enhance its current compensation package. In fact, for a business relying on its members to extract, elaborate, and improve industrial processes, a deep understanding of KPIs and their performance is fundamental.

Four clients of fischer Consulting Italia were selected for specific metric evaluation.

The statistical analysis encompasses different phases: data gathering, manipulation of observations, categorization, correlation analyses, standardization, regression analyses, and time series modeling using ARIMA. These steps were necessary in order to possibly have a statistical base on which we could build an evaluation of the uncertainty in introducing a potential compensation package with a variable part tied to the firm's risk aversion.

In conclusion, this thesis delves into the statistical analysis tool that can be implemented to evaluate the possible impact that consultancy have on their clients KPI, with a vast consideration regarding the potential bias and issue observed in such analysis. Furthermore, it explores the potential of using real life data to improve compensation offer, which is a relevant aspect of a thriving business such as fischer Consulting Italia.

1. CHAPTER 1: Statistical analysis, potential threats, and limits

1.1 Definition of statistical analysis

Statistics is an ancient science, with roots tracing back to the thirteenth century when it was initially employed to compute the number of Arabic words without vowels through the use of permutations and combinations. However, since then, a significant evolution has taken place. The modern term "statistics" originates from the German word "*Statistik*" referring to the description of a state or country. This etymology underscores those statistics involves describing characteristics or information derived from the study of fixed elements, commonly referred to as data. In order to deeply understand the meaning of the word statistic, it is crucial to differentiate between probability and statistics. Probability revolves around predicting the likelihood of future events, whereas statistics delves into the analysis of the frequency of past events. Moreover, statistics is an applied science that predominantly concentrates on real-life observations, while probability represents a more theoretical branch of mathematics, which mostly focuses on elucidating the consequences of mathematical definitions.

Another fundamental question that may arise from the previous definitions is what it is a data. As exacerbated in the book "La nuova intelligenza digitale" of Gaetano Bruno Ronsivalle the structure of data is complex and composed of multiples parts.

As written by the professor, a data can be defined, in a generic manner, as:

"Result of the intentional observation process of a magnitude within a specific time interval and in a particular reference context."

More specifically, the professor elucidated the intricate structure of data by examining five key aspects: object and representation; imagery with filters; data versus information; digital integrity; and, finally, composed contingency.

Firstly, data serves as a representation of what we want to know about, hence it is not the "thing" itself, or as Magritte, the well-known surrealistic painter, draw: "*Ceci n'est pas une pipe*"; secondly a data is an image with filters, since they are shaped by conventional representations and are influenced by various factors. Thirdly, it is crucial to differentiate a data from an information; while the former is a structural component of a specific information, the latter can be considered as the result of a system of data arranged in a peculiar way between them. Consequently, data emerges from a logical and procedural construction whose nature is artificial, and not atemporal since it is inherently defined by the place and the contingencies in which is generated.

Over time, statistics has evolved into a substantial body of work, leading some to argue that it should be recognized as a distinct science, separate from its ancestor, mathematics. Today, statistics permeates nearly every aspect of society, playing a role in industries, tracking outputs, recording birth rates, with its core application in the collection, organization, interpretation, and presentation of analysis of data.

1.2 Objective of statistical analysis

Having grasped the fundamental nature of statistics, we can delve into the concept of statistical analysis. We can define the statistical analysis as an examination that employs statistical tools, techniques, and methods developed in the field of statistics. The primary objective of statistical analysis is to provide analysts with general or specific information about a set of data, which must be more than a singular data point. In essence, statistical analysis enables the connection and correlation of multiple data points or individual observations using defined techniques and methods. These yields summarized information regarding the set, allowing users to comprehend their data as a cohesive whole, rather than focusing on isolated data points.

1.3 Boundaries of data gathering process; origins, accuracy and means of measurement

Before delving into the data gathering process, it is pivotal to clarify the definitions and roles of the terms population and sample. The population encompasses the entirety of items, individuals, or events relevant to a specific question. On the other hand, a sample represents only a subset of the population. We utilize inferential statistics on the sample to gain insights into the probability characteristics of the population. This approach enables us to draw probabilistic conclusions regarding hypothetical data extracted from the population.

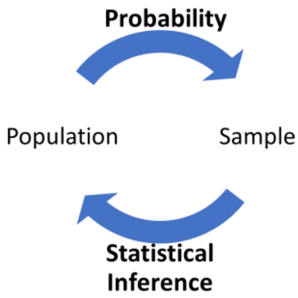

Figure 1: Fundamental relationship between probability and inferential statistics, Figure 1.2 "Probability & Statistics for Engineers & Scientists"

Recognizing this, the initial phase of the data gathering process doesn't involve collecting actual data. Instead, the key question is whether the analyst, with a certain degree of confidence, can effectively address the posed question. Furthermore, it is crucial to assess whether the anticipated answer is both correct and statistically significant. Consequently, the selection of data and the creation of a pertinent sample hold greater significance than the subsequent data analysis or insights. If the data are biased, any analysis conducted is likely to be similarly affected by these biases, leading, probably, to unpredictable outcomes.

This underscores the fundamental importance of selecting a representative sample, with a common assumption being simple random sampling, as it is identified at page 27 of "Probability & Statistics, for Engineers & Scientists". This implies that the probability of selecting any particular sample is equal to that of any other sample of the same size. If not assumed, this principle may result in a failure to accurately represent the overall population by sampling, which implies that the analysis is going to be ineffective in answering the selected question. Nevertheless, it is pivotal to underline that biased sample do indeed exist and in order to address populations which are not heterogenous and that naturally divide into nonoverlapping subgroups that are homogeneous, many sampling techniques have been developed.

To obtain the most reliable data, various factors must be considered, including origins, accuracy, precision, means and methods of measurement, homogeneity in recorded measures, verifiability of data, units of measure, and frequency of measurement. It is crucial to emphasize that all these factors introduce variability external to the data points themselves. Even dough, this variability is not directly influencing the actual value of the data points, it is indirectly impacting the observations' recorded value. Therefore, these factors can be viewed as exerting a unidirectional influence on the recorded data.

Commencing from the beginning, it's important to know where the data comes from – whether it's from a survey, interview, machine, or something else. We need to figure out if a person or a machine is behind this information and if it's gathered from inside or outside the organization. The reason why this matter is because where the data comes from can affect how it's collected and recorded, and it might explain the different patterns or mistakes we find in other factors. In conclusion, this helps us decide if we can rely on the data or not.

Subsequently, we need to evaluate accuracy, which is about how close the recorded data is to the real value we wanted to record. In simple terms, we want the recorded number to be very close to the actual one.

Precision is another important thing to consider. It's about how close different results are to each other. Accuracy is like hitting the bullseye once, while precision is about hitting close to the same spot multiple times. Imagine throwing darts – accuracy is hitting the center, and precision is getting your darts close together. So, when we talk about precision, we're looking at results over time because we want things to be consistent through it. The previous explanation can be summarized in the following image:

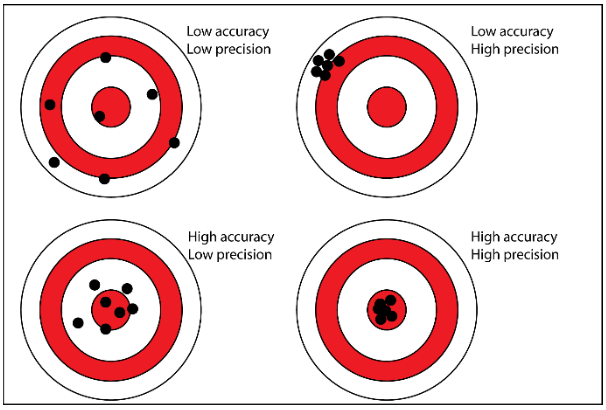

Figure 2: Graphical representation of difference between accuracy and precision from www.antarcticglaciers.org

Precise data collection is essential for drawing accurate conclusions. Hence, evaluating the tools employed in data collection becomes crucial to minimize errors and ensure the faithful representation of the studied phenomenon. Moreover, understanding the means of measurements enables us to grasp potential errors linked to the sensitivity of the recording machine, the frequency of recording errors, the nature of these errors, and whether they can be identified and corrected.

The method of measurement is closely linked to the means of measurement, which can vary from individual checks to sampling methods or experimental designs. Each method has its complexities in assessing recorded values, necessitating an understanding of their implications. However, these intricacies are not explored further in this thesis. Consistency and uniformity in recorded values are also crucial, implying stability across measurements. Factors such as machine switches or human involvement introduce potential differences in accuracy, precision, and error frequency, requiring careful consideration. Additionally, if more people are involved in the registration of the data, there's a risk of misinterpretation, introducing another source of error.

Climate and temperature can also impact the performance of tools, hindering their ability to measure values accurately. In conclusion, for robust and comparable data, these influencing factors should have a high degree of consistency and similarity over time, allowing for a high level of precision.

Verifiability is an underpinning characteristic for a solid data gathering process. In fact, the ability to check and correct erroneously recorded data is essential. Without knowing whether a data point was measured correctly or not, we cannot determine the statistical relevance of our sample. Making computations and analyses on incorrect data can lead to erroneous conclusions.

In today's professional landscape, Key Performance Indicators (KPIs) play a pivotal role in monitoring and evaluating the performance of specific areas, actions, or conditions within various industries and businesses. These KPIs serve to quantify business activities, providing valuable insights into the functioning of operations and often influencing employee compensation structures.

However, in professions where fixed KPIs or objectives must be met within pre-determined timeframes, there exists a potential risk of incentivizing the reporting of inaccurate performance. Therefore, ensuring verifiability becomes essential to validate, rectify, and fortify the data collection process and subsequent analyses reliant on this data.

Another often overlooked aspect is the units of measure. Understanding the units in which each machine or human reports their values might be influential. Comparing grams with kilograms or watts with pounds is not feasible, so having clear information on units of measure is necessary for analyzing the data.

Frequency is another principal information, particularly when and how often a specific observation is recorded. Moreover, when comparing different data points, knowing the recording date is wise, as external factors may influence the performance or value recorded on a particular day, month, or year.

1.4 Bounded knowledge, general problem when dealing only with data

When dealing with sensitive information like personal data, proprietary company details, or national statistics, constraints often arise, whether they be legal, due to confidentiality and security requirements, or related to data minimization, cross-border transfer restrictions, and intellectual proprietary rights. In such instances, analysts may find themselves wrestling with limited data or facing challenges in explaining unexpected results.

In the specific case of the fischer Consulting Italia study, I encountered a unique scenario. I had access to limited data; furthermore, the intensity and impact of the aforementioned factors remain unknown. Similar to an open system, our data points interact externally with these influencing factors, but I lack information about elements in the surroundings that could have affected the recorded values, or the recording, of these data points.

Similarly, to the black box in airplanes, the data may provide insights into what occurred, but without the broader context, or the rest of the plane, assessing the significance and relevance of the results and the analyses becomes more testing.

Another complication arises from the unpredictability of results when dealing with a black box type study. It becomes challenging to anticipate the outcome of the analysis and choose appropriate statistical tools to handle complications or difficulties. Traditional models like Gage R&R, which rely on specific information such as measurement system variation, operator influence, and the system's capability to discriminate between different parts, are not feasible in this context due to the lack of access or knowledge.

In conclusion, while the usefulness of this analysis is certain, it can be assumed, with a high degree of certainty, that the analysis will be incomplete, which may cause the overlook of causes of specific outcomes.

2. CHAPTER 2: Advanced statistical analysis

2.1. Descriptive Statistics

When dealing with datasets, the initial focus often lies on computing descriptive statistics. As the name implies, these statistics are used to describe the characteristics of the data points from which they are derived. Precisely, descriptive statistics can be defined as a set of statistical measures that summarize, organize, and present meaningful information in a concise manner. Even more narrowly, descriptive statistics do not aim to make inferences about an entire population based on a given sample or samples, as is the case in inferential statistics. Instead, they might provide analysts with insights into the trends, patterns, and distributions within the dataset.

For example, descriptive statistics encompass measures of central tendency, such as mean, median, and mode, as well as measures of dispersion, including range, variance, and standard deviation. Additionally, they involve aspects related to the shape of the statistical distribution, such as skewness and kurtosis.

While the classification of what constitutes descriptive statistics remains a subject of debate, some scholars include graphical representation in this category. Graphs and charts that aim to summarize and enhance the understanding of descriptive statistics are considered descriptive, as they visually depict the dataset's characteristics.

2.2. Standardization: Z-Score

In the comparison of different data samples, addressing peculiar characteristics, or visualizing dataset distributions, standardization emerges as a valuable tool. This process involves transforming data points to normalize them, making it easier to interpret and compare datasets. Standardization is a specific technique within the broader category of normalization methods, all aiming to establish a consistent scale across various samples or measurements, thereby enhancing compatibility between different datasets or variables.

Before delving into the advantages and disadvantages of standardization, it is essential to understand how it is computed in this thesis. The widely accepted method for calculating the standard score, or Z-Score, involves determining the number of standard deviations by which an observed data point deviates from the mean. Mathematically, if the population mean (μ) and standard deviation (σ) are known, a raw score (x) is converted into a standard score using the formula:

$$z = \frac{x - \mu}{\sigma}$$

In cases where the population standard deviation or mean is unknown, sample standard deviation and sample mean are used as approximations.

As previously highlighted, standardization offers various benefits, including outlier detection. Outliers, positioned at the extreme ends (positive or negative) of the Z-Score distribution, may indicate data points not representative of the 'normal' observed values. Identifying and discarding these outliers is crucial, especially in developing prediction models to prevent noise.

Another significant application of computing Z-Scores is evaluating the distribution's resemblance to a normal distribution. For example, assessing if approximately 68% of observations fall within one standard deviation, 96% within two, and so on. This analysis informs whether approximating the data distribution to a standard one might lead to substantial errors or is a suitable approach.

Moreover, standardization simplifies model interpretation by translating all variables into standard deviations from the mean, establishing a common level for comparison. In regression analysis, standardization plays a key role in reducing the overall weights that large observations can exert on the model, providing a fairer assessment of each variable's impact.

Within regression studies, standardization addresses multicollinearity issues, where independent variables are highly correlated. By minimizing the impact of different scales on correlation measures, it improves the reliability of the analysis. Additionally, interpreting coefficients becomes more intuitive, as a one-unit change in the standardized value corresponds to a one-standard-deviation movement away from the mean value."

While standardization provides valuable insights, it is not without its limitations. First and foremost, the process results in a loss of the original units of measurement. In statistical analyses, this means that the nuanced impact of a one-unit change in the regressor on the independent variable is obscured by the standardized unit of measure.

Secondly, standardization relies on a bold assumption about the normal distribution of the data. If this assumption is not met, the resulting distribution may be skewed, failing to accurately represent the relative position of data points. Additionally, the reliance on mean and standard deviation as descriptive statistics, aiming to summarize sample characteristics, introduces vulnerability to outliers. These outliers, influencing measures of central tendency, can distort the portrayal of the majority of recorded values.

It's worth noting that when using Z-Scores to identify outliers, the outliers themselves contribute to the computation of the mean and standard deviation, subsequently impacting the Z-Score calculation. This circular relationship can lead to a distorted representation of the data.

In the academic realm, the significance of sample size is well-established. Larger samples generally contribute to more accurate statistics, enhancing the quality of research outcomes. Consequently, the use of standardization in small sample sizes can exacerbate variability within computed Z-Scores. This increased variability poses challenges in analysis and interpretation, underscoring the importance of considering sample size in the application of standardization techniques.

2.3. Linear Regressions

Linear regression stands as one of the cornerstones in statistical analysis, widely employed to model relationships between variables. Initially designed to predict planetary movement, it gains its prominence through the application in the social sciences. This method aims to gauge the linear association between a dependent variable and one or more independent variables, often referred to as regressors.

At its core, linear regression seeks to assess the predictability of dependent variable values based on the corresponding values of associated independent variables. In a simplified scenario, an apt depiction of the relationship between the independent variable Y and an individual regressor x is through the linear equation:

$$Y = \beta_0 + \beta_1 \cdot x$$

Here, β_0 represents the intercept, and β_1 signifies the slope. In an idealized setting, an exact relationship would imply a deterministic link between two variables, without any random elements. However, this picture is often naive and imprecise in the context of real-world variables and analysis.

Consider the demand for oil, for instance. While it may correlate with the quantity demanded, numerous factors, ranging from geopolitical tensions (e.g., October 1973) to global financial crises (e.g., 2009), influence oil prices. Acknowledging such complexities, most cases necessitate the inclusion of a random component, commonly known as the error or random error. This error, ϵ, assumes an expected value of 0 and a variance of σ^2, with the latter often termed residual variance.

The pivotal assumption is that this random error maintains a constant variance, serving as the element preventing our model from transforming into a deterministic relationship. To sum up, with E(ϵ) = 0, at a specific value of x, the dependent variable y is expected to be distributed around the population regression line.

In conclusion, the essence of regression analysis lies in the estimation of the parameters β_0 and β_0, commonly known as regression coefficients. These coefficients play a crucial role in unraveling the intricacies of the relationship between variables.

2.3.1. Least Squares and OLS

Having examined the theoretical foundations of linear regression, let's now explore a specific method for estimating regression coefficients. Among the frequently employed models, Least Squares stands out, aiming to minimize the sum of the squares of the residuals. Residuals, in this context, represent the errors in fitting the estimated model $\hat{y} = b_0 + b_1 \cdot x$.

In practical terms, a residual (denoted as e_i) is the disparity between the actual recorded value and the predicted value of the model: $e_i = y_i - \hat{y_i}$, where $i = 1,2,...,n$ and n signifies the number of observations considered in the model. Essentially, these residuals serve as the empirical counterparts of the random error.

The Ordinary Least Squares (OLS) method operates by differentiating the sum of the squares of each residual, setting the partial derivatives equal to zero, and subsequently rearranging the terms, leading to the derivation of what is known as the normal equations. For a detailed insight into this derivation, refer to page 416 of Chapter 11 in the book "Probability & Statistics for Engineers & Scientists" by Walpole, Myers, Myers, and Ye.

Upon resolving the equations, the determined least square estimates for β_0 and β_1 are as follows:

$$b_1 = \frac{Cov(x,y)}{Var(X)}$$

$$b_0 = \underline{y} - b_1 x$$

This reveals a clear connection between the covariance of x and y and the value of the coefficient of x. As concisely captured by Nick Huntington-Klein in his book, "The Effect": "This is roughly saying, of all the variation in x, how much of it varies along with y?". Essentially, the ratio of the coefficients illustrates how much y changes as x varies, providing insights into their relationship—whether positively, indicating a direct correlation, or negatively, implying an inverse connection—relative to the general tendency of x to change.

2.3.2. Significance Level and R^2

Running a linear regression using the lm method in R provides crucial information, including the significance level and the R-squared (R^2). However, grasping the relevance of these factors and understanding what they measure necessitates taking a step back.

If the error term in $Y_i = \beta_o + \beta_1 x_i + \epsilon_i$ is independent from sample to sample lays the basement for determining the means and variances for the estimators β_o and β_1. As demonstrated by the authors of the previously referenced book, "Probability & Statistics for Engineers & Scientists", the least squares estimators for β_o and β_1 are unbiased. This implies that, on average, the expected values of the estimators align with the true values.

Moving on to the assessment of R^2, we introduce the sum of squares errors (SSE) or residuals. From this, we can compute an unbiased estimator of the population variance, denoted as $s^2 = \frac{SSE}{n-2}$, also known as the mean square error (MSE). Further details on its derivation can be found on pages 421 and 422 of the aforementioned book.

The subsequent step in estimating the regression coefficients involves making inferences about their confidence intervals. For instance, in the case of β_1, according to the theorems 8.4 and 8.5 in the same book, we can compute a T statistic. This statistic follows a t-distribution with $n-2$ degrees of freedom, enabling the construction of an interval with a confidence of $100 \cdot (1-\alpha)\%$ for the coefficient β_1. This process ensures a robust understanding of the precision and reliability of the estimated regression coefficients.

It follows the hypothetical construction of the confidence interval:

$$b_1 - t_{\alpha/2}\frac{s}{\sqrt{Var(x)}} < \beta_1 < b_1 + t_{\alpha/2}\frac{s}{\sqrt{Var(x)}}$$

The final step in assessing the significance level of our results involves hypothesis testing, depending on two fundamental assumptions: the null hypothesis (H_0) posits that $\beta_1 = 0$, contrasting with the alternative hypothesis (H_1) asserting that $\beta_1 \neq 0$. This dichotomy is allowing the evaluation of the t-statistic, which is tabulated, and deducing the corresponding p-value.

Another crucial metric to monitor is R^2, commonly referred to as the coefficient of determination. It elucidates the proportion of variability explained by the model. To achieve this, R^2 considers the sum of squares errors (SSE), representing unexplained variation due to errors in the model, and the total corrected sum of squares (SST), denoted as: $SST = \sum_{i=1}^{n} (y_i - \overline{y_i})^2$, capturing the variation explainable by the model. In essence, R^2 is calculated as $1 - \frac{SSE}{SST}$. For a perfect fit, all residuals must be equal to 0, emphasizing the model's capacity to account for the observed variability.

2.3.3. Residual analysis: Heteroskedasticity and Homoscedasticity

Subsequently to a linear regression analysis, a crucial step involves examining the residuals through a residual analysis. This task is important as the model's accuracy may exhibit variation with changing values of x. One key consideration is whether the variance of the residuals changes at varying x, leading to three potential scenarios: homoscedasticity, heteroskedasticity, or neither.

A dataset is deemed homoscedastic if the variability of the residuals remains consistent across the entire range of the independent variable(s). In simpler terms, it implies that the variance of errors is the same for all levels of the predictors. This assumption is essential for the reliability of standard errors, confidence intervals, and hypothesis tests in linear regression.

Conversely, heteroskedasticity is the opposite of homoscedasticity. A dataset is considered heteroskedastic if the variability of residuals is not constant across different levels of the independent variable(s). In other words, the spread of errors systematically increases or decreases as the predicted values change. Heteroskedasticity violates one of the assumptions of classical linear regression, potentially leading to inefficient estimates of coefficients and biased standard errors.

Graphically, we can distinguish between a homoscedastic and a heteroskedastic dataset through the following visual cues:

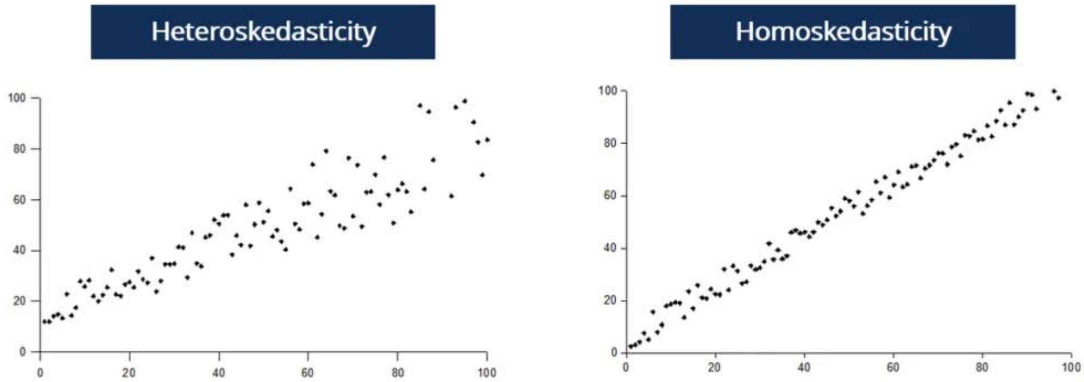

Figure 3: Graphical representation between Heteroskedasticity and Homoskedasticity from Wikipedia

2.3.4. Linear Regression in R

During the elaboration of the regressions results, in our study, the *lm( )* function in R was used. This leverage the QR decomposition, which proved instrumental for handling multiple regressors and executing OLS accurately. However, a detailed exposition of this method lies beyond the scope of this thesis.

Throughout the linear regressive analysis, it has been a recurring scenario to regress performance against the date of observation. Although not optimal, since there is no doubt about the causal relationship between dates and performance, which is none, since a date, or any particular date has no causal effect on the performance achieved itself, since the machines does not function better on Monday, just because it is Monday, but they do so because there might be some other actions or situations that happen solely on Monday and which might be related to a determined influence on the performance. Nevertheless, regression studies correlation, and not causation, hence it remains a valid approach to use *data_obs* (or date of observation) as a regressor.

Dates were stored within the dataset as POSIXct format, which include the specific time and date at which a certain information were recorded, hence in order to make the regressor a progressive integer, it was convert using the function *as.date( )*. This proves beneficial from a theoretical and practical standpoint. This optimization aligns with the expectation that performance would exhibit improvement from the inception of consultant interventions, which was a date.

In our regression model, a key feature often incorporated is the variable *date_obs:after*, serving as an interaction term. This strategic addition enables us to investigate potential variations in the relationship between the *date_obs* variable and the dependent variable, both preceding and succeeding a specific temporal juncture. In our context, the *after* variable identifies whether an observation was recorded prior to or after the initiation of the consultancy project, respectively identified by a value of zero and one.

The *date_obs:after* term, a product of the *date_obs* variable and the binary indicator *after*, assumes the value of 1 for observations occurring subsequent to the commencement of the project and 0 otherwise.

Including this interaction term facilitates the exploration of potential shifts in the impact of *date_obs* on the dependent variable post the specified date. Essentially, it allows for an examination of whether a distinctive relationship exists between the independent variable and the dependent variable in the periods preceding and following this critical temporal marker. The coefficient affiliated with *date_obs:after* becomes a crucial indicator of the magnitude and direction of this observed change, offering valuable insights into the temporal dynamics inherent in the relationship under consideration.

In conclusion, among the three regressions conducted, the most comprehensive one was as follows:

```
m <- lm(performance ~ date_obs + after + date_obs:after, data = data_loc)
```

This snippet of code captures the storage, within variable *m*, of the results of a linear regression characterized by the dependent variable *performance* and all the regressors: *date_obs*, *after*, and the interaction term *date_obs:after*, using the data contained within *data_loc* (local data), which is specific to each individual KPI after filtering:

```
for (kpi in KPI_cat) {
data_loc <- dataset %>%
filter(KPI_id == kpi)
```

The *KPI_cat* variable is defined as a list containing the unique KPI IDs present in the dataset. It is then iterated over to filter the dataset for each specific KPI ID, allowing for separate analyses to be conducted for each individual KPI.

The method used to compute the p-values is particularly intriguing. It leverages the *pt( )* function, which is part of the TDist function and is related to the Student t-distribution. The *pt( )* function is defined as follows: *pt(q, df, ncp, lower.tail = TRUE, log.p = FALSE)*, with our focus exclusively on *q*, *df*, and *lower.tail*, since the other were left at default values.

q represents the vector of quantiles or the values to be tested in our case. Referring to the previous sections, the p-value for the null hypothesis was computed. The p-value was a ratio of values itself, with the numerator stored in the *coeff* lists within the variable m, itself a list of lists. And the denominator as computed by making use of the summary of m, and then taken from the sub-list coefficients. Additionally, the *abs( )* function is employed to ensure that the *coef(m)[2]* value is always positive, if negative at all.

Moving on, *df* signifies the degrees of freedom, computed by subtracting the number of coefficients needed to compute from the number of observations within the local dataset. In our scenario, we needed to compute four coefficients: the intercepts, and then the individual coefficients of *date_obs*, *after*, and *date_obs:after*.

Lastly, *lower.tail*, by default set to *TRUE*, implies a one-tailed test where the confidence interval is concentrated on one side, positive or negative. While useful in certain applications, this wasn't applicable to our case; hence, it was set to *FALSE* to conduct a two-tailed test, effectively splitting the confidence interval into two.

Here's the snippet of code for the computation of the three p-values.

```
p_value_date_obs = 2 * pt(abs(coef(m)[2] / summary(m)$coefficients[6]), length(data_loc$date_obs) - 4,
lower.tail = FALSE),
p_value_after = 2 * pt(abs(coef(m)[3] / summary(m)$coefficients[7]), length(data_loc$date_obs) - 4,
lower.tail = FALSE),
p_value_date_after = 2 * pt(abs(coef(m)[4] / summary(m)$coefficients[8]), length(data_loc$date_obs) -
4, lower.tail = FALSE),
```

2.4. Correlation of variables

In our previous discussions, we operated under the assumption that the observed values of different variables were unrelated to each other. However, in the "real" world, variables often exhibit some form of relationship. Correlation is a statistical measure designed to assess the existence and strength of relationships between two or more variables, whether, or not, they are in a causal relationship.

There are several advantages to evaluating the correlation between variables. Firstly, it provides insights into the potential existence of a causal relationship between two variables. Secondly, correlation analysis is often employed to determine if changes in one variable can predict changes in another. Thirdly, in research, correlation helps us gauge the significance of statistical results and discern whether an observed association is statistically meaningful or merely due to chance.

The primary reason for prioritizing the study of correlation is to prevent the misinterpretation of causal relationships between variables. It is crucial to recognize that correlation does not imply causation. Frequently, there may be a strong correlation between two variables, but causation may be influenced by a third variable. A classic example is the correlation between ice cream sales and the number of drowning incidents, where the shared correlation may be attributed to a common factor—weather.

Typically, the type of correlation tested is linear correlation, which can be assessed using the Pearson product-moment correlation, or simply Pearson correlation. This method aims to measure the correlation coefficient, r, by fitting a line of best fit through the data of two variables. The Pearson coefficient can take values within the range of -1 to 1. A value of zero indicates no association between the variables, -1 indicates a negative correlation (one variable tends to decrease as the other increases), and +1 indicates a positive correlation (both variables move in the same direction).

The calculation of the Pearson coefficient is intuitive, involving the ratio of covariance to the product of the standard deviations of the two individual variables, as expressed by the formula:

$$r = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}, \text{ if } \sigma_X \sigma_Y > 0$$

This formula essentially captures the tendency of the two variables (X and Y) to change together in relation to their individual standard deviations. For a comprehensive understanding of the calculation procedures, further details can be found in subsection 11.12 of the book "Probability & Statistics for Engineers & Scientists".

2.4.1. Correlation test in R

In the R program, it was chosen to make use of the *cor.test( )* function, in order to address the correlation between two specific variables: *diff_since_beg*, and *performance*. Specifically, *cor.test( )* is designed for computing one of either three method of correlation: Pearson, Kendall, or Spearman. It was decided to use the Pearson correlation coefficient for paired samples, as in our case. Paired samples involve items or individuals appearing consistently across all compared samples. In contrast, unpaired samples may have instances where items or individuals do not appear in every sample. It was decided to use the paired one because we can expect that for each observation in one variable, like it might be performance, there is an associated value in the other variable. More generally, for each data point in one variable, there is a corresponding data point in the other variable that refers to the same individual, object, or event.

The code use for doing this is the following:

```
correlation_results<-list()
for (kpi in KPI_cat){
  data_loc<-dataset%>%
    filter(kpi==KPI_id)
  correlation_res<-cor.test(as.numeric(data_loc$diff_since_beg),data_loc$performance)
  cat("KPI:",kpi,":\n")
  print(correlation_res)
  correlation_results[[kpi]]<-correlation_res
}
```

In this excerpt, a list was used to aggregate the results of the correlation analysis between two variables: the performance metrics and the difference in days since, or from, the inception of the project. Hence, this variable will yield negative values for data recorded before the project's commencement and positive values for data captured thereafter. Moreover, the variable diff_since_beg assumes numerical form, as it represents the disparity between two variables.

2.5. LOESS

LOESS, or Locally Weighted Scatterplot Smoothing, is a non-parametric regression technique designed for analyzing and visualizing data relationships. Differing from global regression methods that assume a constant relationship across the entire dataset, LOESS adapts to local variations, making it particularly effective for capturing complex, non-linear patterns.

To construct a predictive model with LOESS, the initial step involves collecting data, specifically values for the predictor variable (x) and the response variable (y). The term "non-parametric" denotes that LOESS utilizes the entire dataset for estimation rather than relying on predefined parameters, making use of all sampled values of y.

The introduction of a tunable parameter, k (window size), is the next step. Larger values of k introduce bias, while smaller values increase variability. The algorithm identifies the k nearest neighbors of a given x' through Euclidean distance calculation. These distances form a set D, transformed into an ordered set W of weights using a tri-cubic weighting function. Normalization of these weights ensures that larger distances result in lower weights, emphasizing the local nature of LOESS.

The process involves converting the set of distances (D) into weights (W), using a specialized function that assigns importance to each neighbor based on its distance to x'. These weights play a crucial role in the subsequent linear regression process.

The final step is the computation of a weighted linear regression for each target point (x'). LOESS establishes a linear regression model, utilizing the k nearest neighbors and their associated weights to predict the corresponding output (y'). Typically, LOESS models low-dimensional polynomials, often a line or a quadratic.

In summary, LOESS produces a series of local weighted regressions, resulting in a flexible and adaptive model that excels at capturing intricate relationships within the data. The algorithm's localized emphasis is visually represented in the outcome, creating a composition of locally fitted regression segments.

Here is an example:

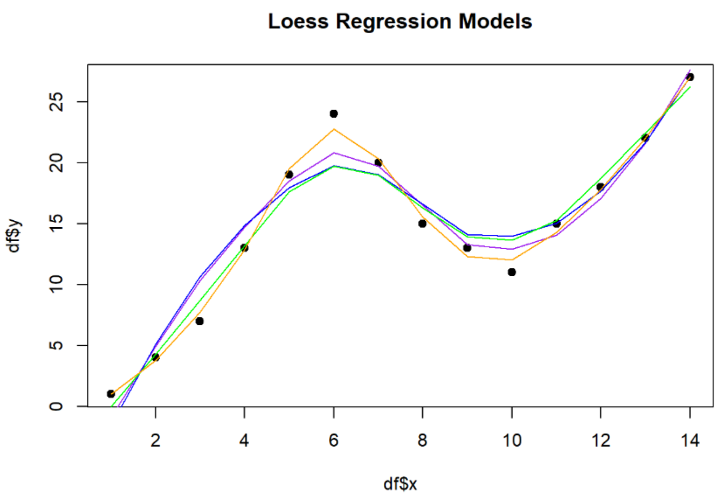

Figure 4: Loess regression models for different spans from r-bloggers.com

2.5.1. LOESS in R

In our software it was use the *loess( )* function, which is part of the package of statistics. The following represent the snippet of code:

```
for (kpi in KPI_cat) {
  for (var in c(0, 1)) {
    data_in_use <- dataset %>%
      filter(KPI_id == kpi & after == var) %>%
      select(date_obs, performance, after) %>%
      mutate(KPI_area = as.factor(KPI_area),
             date_obs = as.integer(date_obs))
    if(nrow(data_in_use)>0){
      loess_model <- loess(performance ~ date_obs, data_in_use)
      loess_summary <- summary(loess_model)
    }
```

Similarly, to the approach taken for the linear regression analysis, the dataset underwent a filter before conducting the loess regression; and then there was a subsequent storage of results in a vector of lists. However, unlike the previous procedure, only a subset of the whole dataset was retained. Furthermore, the variable *date_obs* were transformed into numerical values, along the transformation of *KPI_area* into a factorial, which in this case will contains only one value, since the data has been previously filtered by an individual KPI and period. Additionally, a check was conducted to ensure that the number of rows in the *data_in_use*, or the current data, exceeded zero. This precautionary measure was implemented to address instances, such as with KPI4, where data for a particular KPI may have been recorded solely before or after certain time points.

2.6. ARIMA

ARIMA, which stands for Autoregressive Integrated Moving Average, is a statistical model which is often used to better comprehend the data and their movement trend, but also to do forecasting of a time series. The model is fundamentally composed of three key parts: autoregressive, integrated and moving average.

2.6.1. Integrated (I)

The integrated component of the ARIMA model involves a series of differencing steps aimed at eliminating the trend and providing the series stationary. This last step is fundamental to simplify the modeling process. The rationale for seeking stationarity lies in the ARIMA model's assumption that the underlying data is stationary, as the statistical properties such as mean and variance should remain constant over time.

Differencing is the computation of the differences between consecutive observations, mathematically expressed as $Y_t' = Y_t - Y_{t-1}$, where the Y_t is the original time series, and Y_t' is the differenced time series.

A critical aspect of the ARIMA model is the determination of the order of integration, denoted as *d* in the *ARIMA (p, d, q)* model. *d* represents the number of differencing steps required to achieve stationarity. For example, if only one differencing step is needed, the order of integration is equal to one.

The assessment of stationarity is conducted through statistical tests such as the Augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The ADF test's null hypothesis is that the time series data has a unit root, indicating non-stationarity. A p-value below the significance level (typically 5%) leads to the rejection of the null hypothesis, suggesting stationarity. On the other hand, the KPSS test examines the null hypothesis that the data is stationary around a deterministic trend. A p-value above the significance level indicates non-rejection of the null hypothesis, suggesting stationarity.

Further details on the ADF and KPSS tests can be found in the Journal of American Statistical Association, Vol. 74, pp. 427-431, and the Journal of Economics, 54(1-3), pp. 159-178, respectively.

The final step involves backward transformation, which means converting the differenced time series back to its original scale. This transformation is essential for utilizing the newly formed time series in the subsequent autoregressive (AR) and moving average (MA) processes of the model.

2.6.2. Autoregressive (AR)

The autoregressive (AR) component of a time series entails leveraging past values to predict its current state. In an AR(p) model, the current value is modeled as a linear combination of the previous p values, where p represents the number of lagged observations included in the model.

For instance, an AR(1) model is represented as: $Y_t = \phi_1 Y_{t-1} + \epsilon_t$, where ϕ_1 is the autoregressive parameter for lag 1, Y_t is the current value, Y_{t-1} is the value at the previous time step, and ϵ_t is the error term.

A pivotal activity is to select the number of lagged observations correctly. This number is denoted by the parameter p in an *ARIMA (p, d, q)* model, and is chosen through a process known as model selection. There are different ways to efficiently select the number of lagged values, we will look at two of them.

Firstly, we do a visual inspection by plotting the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the time series data. ACF shows the correlation between the series and its lagged values, while PACF shows the correlation between the series and its lagged values with the effects of intervening lags removed. By lags removed is meant that it measures the direct relationship between the series at a certain lag and the series at the current time point, excluding the influence of the lags in between. Peaks in the ACF and PACF plots indicate potential values for 'p'.

Secondly, model selection criteria like the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) are employed. These criteria balance model fit with complexity. Lower AIC or BIC values indicate better-fitting models. Experimenting with different 'p' values, fitting ARIMA models, and choosing the one with the lowest AIC or BIC enhances model precision.

The methods in which AIC and BIC operate, though similar, have peculiarities. AIC tends to favor more complex models, especially with larger sample sizes, while BIC leans toward simpler models, particularly with smaller sample sizes. Both criteria don't promise a validated model but offer a means of comparison, as evident in their formulas: AIC $-2 \cdot log(L) + 2 \cdot k$ and BIC $-2 \cdot log(L) + k \cdot log(n)$, where L is the likelihood of observing a certain value under a model, k is the number of parameters, and n is the sample size.

Given this, the optimal choice of p may vary based on specific time series characteristics. Experimentation with different values, visualization, and application of statistical criteria contribute to refining the predictive model. Once p is determined, the linear combination of previous p values is computed, represented as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \epsilon_t$$

Here, ϕ_x denotes the coefficients or strength and direction of each lagged value, estimated through ordinary least squares (OLS) processes. For detailed explanations, refer to "Time Series Analysis and Its Applications: With R Examples" by Shumway and Stoffer (2010).

2.6.3. Moving Average (MA)

The third and final component of the ARIMA model is the moving average (MA), which, as the name implies, calculates different averages for various subsets of the full dataset. The purpose of this approach is to capture short-term dependencies in the time series data that cannot be explained by past observed values but are attributed to random shocks. In a sense, the MA acts alongside the integrated component (I) to achieve stationarity, enhance model predictability, and reduce its complexity.

The previous random shocks are typically referred to as white noise, which consist of a sequence of uncorrelated and identically distributed random variables with a mean of zero and constant variance.

The MA model computes the q component, indicating the number of past white noise terms included in the model to enhance its predictions accuracy.

The mathematical representation of the MA model is given by the formula:

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q}$$

Here, Y_t is the observed value at time t, μ is the mean of the time series, ϵ_t is the white noise at time t, and θ_1, θ_2, …, θ_q are the parameters of the MA model.

These parameters are estimated using a method called maximum likelihood estimation (MLE), which measures how well the model explains the observed data. MLE involves taking the derivative of the likelihood function to maximize it. The likelihood function provides the probability of observing the given data under different parameter values. It's crucial to note that a probability density function conveys the probability of observing the data given the defined distribution parameters, and these parameters must be known. However, the likelihood function expresses the likelihood of parameter values given the observed data, assuming the parameters are unknown.

Working out the distribution function can be challenging, leading to the frequent use of the log-likelihood function. For a more in-depth understanding of the proof and function of this model, Chapter 5 of "Mathematical Statistics: An Introduction to Likelihood-Based Inference" by Richard J. Rossi provides comprehensive information.

2.6.4. ACF and PACF functions

In subsection 2.6.2, we discussed the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), which are essential tools in time series analysis.

The ACF assesses the correlation between a series and its lagged values, revealing the level of similarity between observations at different time points within the same series. Mathematically, the ACF at lag k is computed as the correlation coefficient between the original series and its lagged version at lag k.

In the ACF plot we find the lag on the horizontal axis and correlation coefficient on the vertical axis. The ACF plot helps us in identifying repeating patterns or structures in the data, possibly indicating seasonal or cyclical behavior. Like Pearson's correlation, ACF values range between -1 and +1, signifying negative perfect direct correlation and perfect direct correlation, respectively.

When analyzing the ACF plot, significant peaks highlight correlations between the series and its lagged values, potentially indicating underlying patterns or dependencies. Decay rate, observed as the decrease in ACF values with increasing lag, provides information into series persistence, with slow decay suggesting a long memory process and pattern repetition, while rapid decay suggests short memory or randomness.

The PACF complements the ACF by measuring the correlation between a series and its lagged values while removing the effects of intervening observations, hence capturing direct correlations between observations at different lags.

Similarly to the ACF plot, the PACF plot is useful in identifying the order of the autoregressive (AR) process in time series data. Some peculiar features include initial spikes, signifying strong direct correlations between the series and its lagged values, which suggest the presence of an autoregressive process. The cut-off, or the lag at which PACF values drop to zero or become non-significant, provides an indication of the order of the AR process, helping in determining the appropriate number of autoregressive terms in ARIMA modeling.

To interpret the graphs effectively, we use the following table:

	AR (p)	**MA (q)**	**ARMA (p, q)**
ACF	Tails off (*Geometric decay*)	Significant at each lag *q* / Cuts off after lag *q*	Tails off (*Geometric decay*)
PACF	Significant at each lag *p* / Cuts off after lag *p*	Tails off (*Geometric decay*)	Tails off (*Geometric decay*)

Table 1: ACF and PACF interpretation table for AR and MA from towardsdatascience.com

Although it directly relates to ARMA, understanding the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) is crucial as ARIMA models are built upon ARMA.

For example, consider the following interpretation:

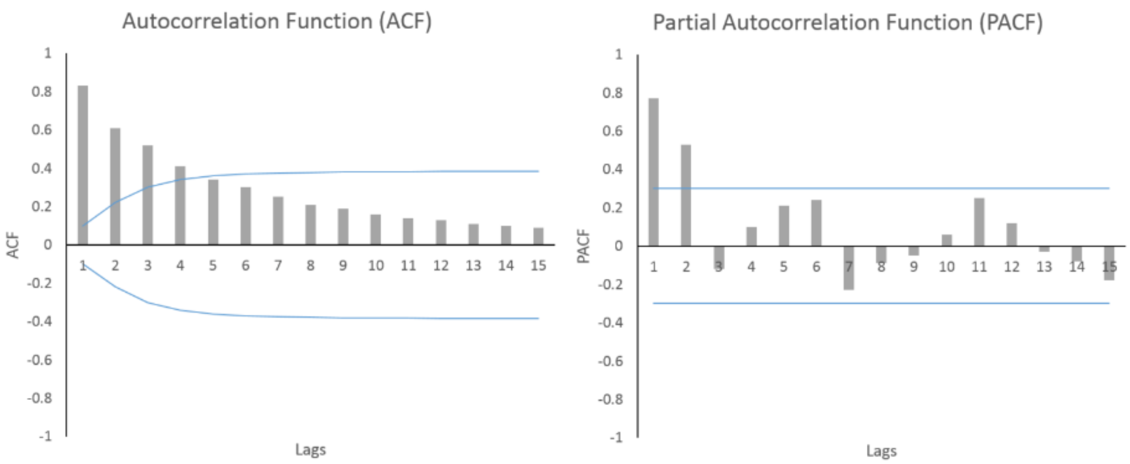

Figure 5: Example of ACF and PACF plots for AR(2) scenario from spureconomics.com

In this case, the ACF exhibits a tails-off pattern, while the PACF reveals only two significant spikes above the blue line before trailing off. Referring to the interpretation table, we can conclude that we are dealing with an AR(2) scenario. This indicates that the current value of the time series depends on the two previous values, expressed mathematically as $y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t$ where c represents the mean of the time series and ϕ_1 and ϕ_2 are the coefficients derived from the PACF.

Conversely, the subsequent graph depicts a situation indicative of MA(2). Here, the PACF displays a tails-off pattern, while the ACF exhibits only two significant spikes. Referring to the interpretation table, we confirm a MA(2) model, mathematically expressed as $y_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$ where c represents the mean of the time series and θ_1 and θ_2 are the coefficients derived from the ACF.

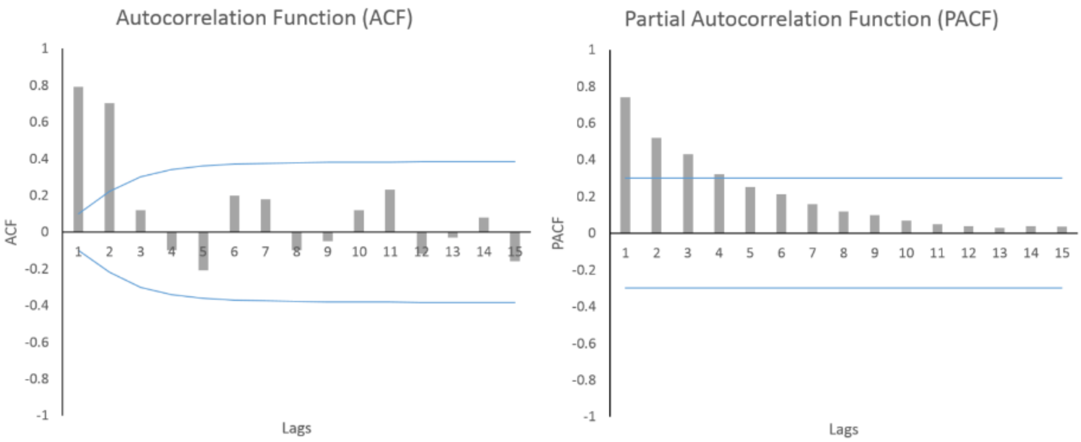

Figure 6: Example of ACF and PACF plots for AR(2) scenario from spureconomics.com

Further interpretations can be made analogously.

2.6.5. auto.arima( ) function in R

During the analysis of the observed data, the *auto.arima( )* function, an integral part of the forecast package, was used. This function employs a variation of the Hyndman-Khandakar algorithm, exploiting the models explained earlier to derive an *ARIMA (p, d, q)* model. Specifically, *auto.arima( )* utilizes the log-likelihood function for the MA process and AIC and BIC for the AR process.

To automatically determine the best-fitting model, the *auto.arima( )* function follows these steps:

1. The number of differences is determined through the KPSS test and is: $0 \leq d \leq 2$.

2. The values of p and q are subsequently chosen by minimizing the AICc after having differentiated the data d times. Since, taking all possible combinations of the three parameters: p, d, and q, would be excessively time-consuming the algorithm takes a stepwise approach. Subsequently, values for p and q are chosen by minimizing the AICc after differentiating the data d times. Due to infeasibility in testing all possible combinations, to manage computational complexity, the algorithm employs a stepwise approach.
    a. Four models are fitted:
        i. ARIMA (0, d, 0)
        ii. ARIMA (2, d, 2)
        iii. ARIMA (1, d, 0)
        iv. ARIMA (0, d, 1)

       A constant is included unless $d = 2$ and if $d \leq 1$ an additional model is also fitted: *ARIMA (0, d, 0)* without a constant.
    b. The model with the smallest AIC value is set to be the “momentary model.”
    c. Other versions of the “momentary model” are considered, by varying p and/or q by ± 1, and by including or excluding the constant. The best one of these variations will become the new momentary model.
    d. We repeat step c until no lower AIC can be achieved.

3. CHAPTER 3: The real study

3.1. fischer Consulting Italia: overview

fischer Consulting Italia Sas is a dynamic business consultancy dedicated to enhancing organizational efficiency and refining processes through a robust commitment to *lean* management principles. Established in 2017, the company has experienced consistent growth, forging partnerships with industry giants in the manufacturing and industrial sectors. Notable clients in its portfolio include GlaxoSmithKline (GSK), Jungheinrich, IRSAP, and the esteemed fischer Group.

In a nutshell, with lean management principles we referred to a management philosophy that is centered on maximizing value while minimizing waste. The core principles, as identify in "Lean Thinking: banish waste and create wealth in your corporation" by the authors James P. Womack and Daniel I. Jones, are five: value, which consist in understanding what the customer is willing to pay for; mapping value-stream, which analyze and map out every process related step that delivers the product or service at hand, and at the same time it identifies where value is added and where waste occurs; create flow, which is the capabilities of the process to ensure a smooth flow of work by eliminating bottlenecks and productive constraints; pull, which coincide with the principle to let customer demand pull the production or services, rather than pushing the production of good or services based on forecasts or schedule. Continuous improvement, is the last principle and it aims to build a model within the company and its employees to continuously strive to improve processes, products, and services by eliminating the *muda (*Japanese term for waste), increasing efficiency, and enhancing quality.

3.1.1. fischer Group

The fischer Group is a multinational conglomerate and a global leader in fixing systems. It also operates across diverse business sectors. With 11 productive sites and 50 subsidiaries across 38 countries, the group boasts a distribution network reaching over 120 countries. In 2022, the group's revenues surpassed €1.1 billion, supported by a workforce of over 5500 employees and a repository of 1500 active patents.

Comprising seven distinct business divisions, the fischer Group showcases its versatility and innovation:

- fischer Fixing Systems: specializing in the production of fixing solutions encompassing nylon, steel, and chemical components.

- fischer Automotive: focused on crafting cup holders and multifunctional consoles, primarily providing to German automakers.
- LNT Automation: dedicated to the development, production, and distribution of tactile interface systems.
- fischer Innomation: positioned as a partner for intricate assembly automation solutions.
- fischer Innovative Molds: specialized in the creation of high-productivity precision molds.
- fischer Consulting: empowering companies of fischer's expertise to enhance processes and minimize production waste.
- fischertechnik: producing intelligent construction kits that facilitate detailed simulations of machinery and industrial structures.

Together, these divisions underline the fischer Group's commitment to innovation, quality, and comprehensive solutions across a spectrum of industries.

3.1.2. fischer Consulting

fischer Consulting Italia emerged after the matured experience of its German counterpart, fischer Consulting GmbH, since 2006. Originally conceived to support fischer's internal divisions in navigating intricate improvement endeavors, the mission has expanded to include exporting this knowledge to external clients, suppliers, and partners.

In the pursuit of continual improvement, Professor Fischer, inspired by a transformative journey to Japan in 2002, where he visited a Toyota production site, crafted a model of the 'fischer Production System' that subsequently became the fPS, or fischer Process System. This initiative, inspired by the renowned Toyota Production System (TPS), marked a revolutionary shift in managerial style within Fischerwerke Gmbh & Co., permeating the entire company with the lean thinking.

The core of fischer Consulting's mission is to guide companies through the transformation of their entire value chain, both in production-logistics and offices, enhancing competitiveness and elevating customer satisfaction. The methodology employed by fischer Consulting is system-driven, focusing on defining objectives, improving performance, and achieving operational excellence. This involves fostering a new way of thinking that emphasizes waste elimination and functional problem resolution.

Operating across five key areas, fischer Consulting Italia specializes in:

- Process improvement: involving the design, implementation, development, and augmentation of direct and indirect processes. This includes waste reduction, optimization of end-to-end value chains, cross-functional improvement of flows, and effective planning, scheduling, and control of production.

- Organizational analysis: conducting health checks on organizations, building pull customer-oriented organizational models and strategic planning.
- Logistics: creating logistics networks, developing material flow models, supplier management, and designing layouts for new facilities and re-layouts.
- People development: focusing on qualifying and boosting the skills of employees through corporate academies, experiential approaches, coaching, and benchmark tours in fischer Italia.
- Support for information systems: providing assistance with ERP systems, Manufacturing Execution Systems (MES), Warehouse Management Systems (WMS), and Business Intelligence (BI) systems.

In essence, fischer Consulting Italia is dedicated to assists companies into a new era of efficiency, strategic growth, and sustainable excellence across their entire operational landscape.

3.1.3. fischer Consulting philosophy

As previously mentioned, fischer Consulting operates in tandem with its primary parent company. In the case of fischer Consulting Italia it is strategically positioned near the Italian production sites of fischer Italia. The latter stands as one of the main global production sites, specializing in plastic fixing systems and providing supports for solar panels. Both Italian entities align with the group philosophy summarized in the fPS, as represented below:

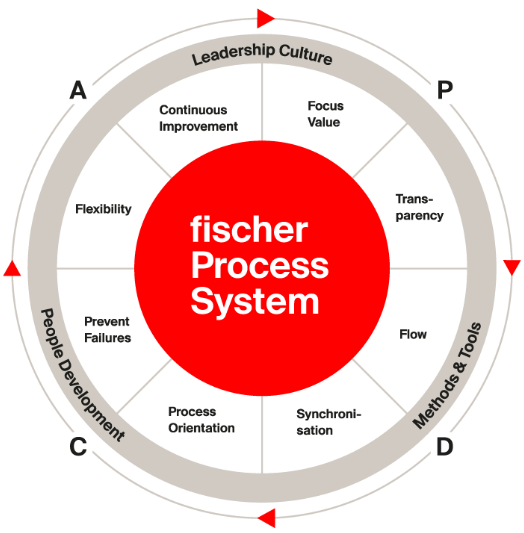

Figure 7: Graphical representation of the fPS (fischer Process System) from fischerconsulting.it

The fPS, structured around eight core principles, drives the operational ethos of the companies:

- Focus on value: prioritizing activities that contribute to customer value while eliminating non-value-adding tasks.
- Transparency: ensure clear visibility in material and information flow, employing Key Performance Indicators (KPIs) for objective measurement.
- Flow: optimize processes to facilitate an ideal continuous flow of materials and information, enhancing speed and visibility.
- Synchronization: align processes with customer demand to prevent overproduction and delays.
- Process orientation: analyze and map processes to enhance communication and coordination between departments.
- Error prevention: implement stable, standardized processes to minimize errors and waste.
- Flexibility: adapt production processes to meet fluctuating market demands.
- Continuous improvement: cultivate a culture of daily, low-cost improvements, establishing new standards for ongoing enhancement.

Effectively implementing the fPS necessitates the utilization of a systematic problem-solving approach, specifically the PDCA or Deming cycle. This method empowers operators to define and assess problems or deviations from the objective, enabling informed decisions to increase the process.

In conclusion, the fPS enables cost control, fosters a commitment to continuous improvement, and mirrors the innovative spirit of its founder, Arthur Fischer, whose legacy of innovation and the creation still leaves nowadays.

3.2. Motivations of the study

The motivation behind fischer Consulting Italia's commitment to studying the impact of their consultancies on their client's performance derive from a genuine desire and curiosity to reveal the existence, or inexistence, of a statistical relationship between performance and consulting interventions.

This analytical approach serves some key objectives:

1. Verify change in average KPI performance:

    a. Tracing trends before and after the commencement of the consultancy projects, and their statistical significance.
2. Understanding relationship and relevance:
    a. Gaining insights into the statistical connection between consultancy efforts and improvements in client performance.
3. Building a track record for clients:
    a. Establishing solid datasets that can be shown to both prospective and existing clients.
    b. Providing customers with a specific numerical representation might instill confidence in the consulting firm, facilitating clients acquisition.
    c. Enhancing the effectiveness of project presentations with concrete proofs of past successes.
4. Improving compensation package:
    a. Exploring the possibility of partially moving away from a compensation model based on expected days of work to one tied to the performance achieved post-consultancy.
    b. This approach should align the interests of fischer Consulting Italia with the measurable success of their clients, encouraging a mutually beneficial relationship.
5. Improving market competitiveness:
    a. Leveraging the gained insights and track record to enhance the overall market competitiveness of fischer Consulting Italia.

In practical terms, these information can be utilized as a dynamic track record shared with clients, contributing to a greater level of trust and comfort in engaging with a consulting firm. Additionally, by incorporating these statistics into SAL (Stato Avanzamento Lavori) or "Work Progress Status" presentations, fischer Consulting Italia can manage and take corrective actions, if needed, based on real-time performance data. This approach not only reflects a commitment to continuous improvement, but it is also underpinning to the adaptability and responsiveness of the consulting activities to client needs.

3.2.1. Economical perspective on performance related compensation package

From an economic standpoint, a compensation structure tied to performance is identified as the most efficient outcome. In the precedent case, the remuneration is linked directly to the concrete achievements that a consultant can deliver. As an entrepreneur or economist, the belief that the value of a consultant's advice or activities is proportional with, or it will exceed, the price paid is fundamental. Without such an assumption, the viability of a project can be questioned.

However, theory and reality often differ from one another. Describing the intricacies of economic relationships requires accounting for an enormous number of factors, a task that was impractical until the advent of computers, over 40 years ago. Even in the present day, identifying these factors and assessing their impact on overall performance remains a challenging effort. Consequently, the early architects of economic theory opted for simplifications in their models, as to favor practicality.

Additionally, the negotiation is frequently affected by informational asymmetry. As outlined in Akerlof and George's work on "The Market for 'Lemons," wherein quality uncertainty can destabilize a market, our reality regularly involves scenarios where the client possesses more information compared to the consulting firm. This knowledge imbalance, due to the client's access to internal firm data and constant interaction with the company's internal operations, allows the buyer additional negotiation leverage, which can potentially distort the accurate monetary evaluation of consultancy services.

In response to this, consultancy firms, including fischer Consulting, implement a crucial step known as a preliminary overview. This approach consists of conducting a vast assessment of the company's state, before the signing of the contract. Because, through site tours and interviews with management, the consultants are able to identify critical aspects, or activities, in need of intervention, and clarify their interconnections with other company's operations. By undertaking this measure, consultancy firms can mitigate the impact of information asymmetry and increase the accuracy in the appraisal of the true value of their services.

3.3. Limiting factors of the consultant

In this study various challenges have arisen, primarily revolving around the availability of data. For instance, while companies are generally cooperative in sharing essential data as to

achieve specific targets or projects, they are reluctant, in diffusing additional information, especially concerning employee details and working hours. This last one, is rooted, sometimes, also in the rightful law protection of personal data and information.

The complexities arising from these limitations imply that while the study remains valuable for understanding implications, it may result as incomplete or incapable of offering a satisfactory explanation for specific performance outcomes.

Subsequently, we will delve into the key factors that may reduce the consultant's capacity to exert influence on Key Performance Indicators (KPI) performance.

3.3.1. Inconsistent presence

A critical aspect in the consultancy process, particularly within fischer Consulting Italia's practice, is the sporadic nature of consultants' presence. This infrequent engagement constitutes challenges as it obstructs the consultant's ability to oversee day-to-day operations, and judge the practical application of their advice, as to ensure its correct implementation.

The significance of recurrent oversight becomes apparent when considering the potential consequences of advice being unutilized or incorrectly applied. Failure to apply recommended changes might lead to unaffected company operations. Moreover, incorrect application of the given advice can result in unintended outcomes, potentially worsening the existing performance rather than fostering the improvement envisioned by agreement between the consulting firm and the client.

While the periodicity of consultant presence varies across companies, fischer Consulting Italia aims to maintain a presence of typically one to two days per week, depending on the type of activity done. This attitude allows for the enactment of corrective measures but falls short of allowing daily checks. Therefore, establishing continuing and intra-weekly communication with clients becomes paramount.

To address this challenge, fischer Consulting Italia emphasizes the importance of having a robust communication structure. This includes the establishment of a steering committee and a dedicated project team; this last one under the direction of a project manager. These entities need to ensure the alignment and commitment of the member to the project goals, and the autonomous execution of concorded activities. This collaborative method should bridge the gap created by infrequent physical presence, making it easier to achieve the settled targets.

3.3.2. Passive Influence

An essential aspect to consider revolves around the type of influence that a consultant, particularly in the context of fischer Consulting, can have, which is predominantly passive. Unlike businesses which are based on renting out personnel (body rental) or offering interim management services, fischer Consulting operates differently. Given the nature of their presence, the consultant's influence is primarily manifested through passive means by offering advice and persuading those responsible for implementing these solutions.

In contrast to an internal manager with the authority to reallocate and reorganize resources, external consultants lack these powers. Due to this distinction, part of the consultant's effectiveness, and by mirroring the project's success, hinges on good communication, collaboration, and the willingness of internal stakeholders to act upon the provided instructions.

3.3.3. Non-exhaustive knowledge of specific processes and people

An additional challenge stems from the intrinsic limitation of the consultant of the internal processes and the specific role and personality of individuals within a company. While well suited at identifying problems and devising tailored solutions, the consultant's knowledge might not include the sophisticated details of how each internal process works or affects other ones. But this gap consists of a trade-off, as the consultant's congenital value lies in possession of a diverse and varied knowledge of various companies functioning.

On a broader scale, effective consultancy must include the engagement of the process owners, or those individuals that are responsible for executing specific activities. Since companies are composed of individuals working towards common goals, any resistance by staff members of a certain project can impede progress towards the desired objective. It follows that some of the difficulties lie in the consultant's inability to identify each employee's personality and opinion regarding the project, since fischer Consulting often collaborates with firms with more than over 150 collaborators, and the consultants can get in touch only with a fraction of them.

A distinguishing quality of a consultant then is the ability to transfer knowledge and expertise to an employee, empowering them to independently carry out the demanded activities. This transfer of skills ensures sustainability beyond the consultant's presence, contributing to the long-term success of the applied solutions.

3.3.4. Exposure of KPIs to external factors

One pivotal factor with a profound impact on the correlation between consultancy advice performance and Key Performance Indicators (KPIs) is the multitude of external components that can influence KPI metrics and inducing fluctuations in KPI performance over time.

Among all the elements capable of having an impact, demand is a good example. A surge in product demand may prompt the company to increase production or expand its current workforce, while shrinking demand could lead to the opposite effect. Such variations can impact tracked KPIs like output or productivity. For instance, the additional workforce may not have the same marginal contribution as the permanent employees to productivity. Furthermore, changes in demand may affect warehouse saturation levels, raw material costs, production costs, revenues, and even profitability of the business due to rising labor costs in overtime scenarios.

The demand is just one of the myriads of factors that can influence KPIs. Consider the situation in which not only demand is changing, but also supply is doing the same. Furthermore, national holidays, periods of company closure, widespread illnesses, seasonal patterns in supply and demand, governmental incentives, new regulations governing workers' conditions, emissions, or health standards, legal issues, new competitors, industry-specific market conditions, rising interest rates, and company acquisitions or sales, might and, probably do, have an impact of the performance of the company, which is summarized in its KPIs.

As it is now clear, the challenge lies in the volume of factors that should be considered and tracked to fully grasp the performance dynamics. However, in reality not all information are readily available or shareable, either due to its sensitive nature, or its confidentiality prevents the data from being disseminated.

Coupled with the complexities highlighted in Chapter 1, Subsection 1.3 regarding the challenges in the data gathering processes the difficulty in executing these analyses becomes apparent, as is the ability to discern the impact of the consultant from the external parts.

3.3.5. Undetermined economic impact

A critical underpinning in determining the value of consultancy, as detailed in Subsection 3.2.1, lies in the obstacle of quantifying its economic impact. Consultants frequently lack access to internal information revealing the profitability of specific investments or the efficiency gains achieved by current equipment improvement operations. This information hole is a persisting problem for consulting firms in precisely determining the value of their services.

The intricacy lies in the strategic aversion of both consultants and clients to disclose insights about potential outcomes. Sharing such information could potentially reveal an overvaluation or undervaluation of consultancy services. Simultaneously, it places the firm in a vulnerable position during negotiations because it may limit negotiation leverage.

The interaction between a consulting firm and a client thus resembles a "black box" negotiation. Both parties are aware of the inputs and outputs, and while this may yield the desired outcomes, the underlying mechanisms remain unknown. This, once again, underscores the complex nature of consultancy value assessment. The difficulties lie in navigating these strategic positions and creating trust while maintaining confidentiality, as to fulfill the need of a black box negotiation.

3.4. Data acquisition and categorization

The data utilized in this analysis were sourced from both ongoing and concluded projects involving current and past clients of fischer Consulting Italia. The possession of these data by fischer Consulting Italia is due to the expressed keen interest of its clients in the development and implementation of digital tools and BI visualization tools, which, in order to be tested, require a similar and significant sample. Additionally, some data were collected to assess the empirical performance of internal operations, aiming to gauge potential improvements in production, quality, and efficiency.

It is crucial to note that these clients willingly shared the data to align with their agreed-upon goals, as outlined in their collaboration agreement with fischer Consulting Italia. No Non-Disclosure Agreements (NDAs) were signed regarding the information or its disclosure, emphasizing the importance of maintaining confidentiality and trust awarded to fischer Consulting Italia by its clients. While a brief introduction to each tracked company and an overview of the monitored Key Performance Indicators (KPIs) will be provided, it is essential to recognize that the data may not be recent and may not accurately reflect the current conditions or operations of these companies.

Following the data acquisition from various sources, primarily Excel files, and subsequent aggregation into a unified structure, each tracked KPI was categorized within a framework. This categorization serves multiple purposes, including differentiating the potential effectiveness of consultants' interventions based on the specific problem area addressed.

Moreover, the categorization aids in evaluating potential heterogeneity or homogeneity in the effectiveness of interventions within a particular organizational area. For example, variations in the quantity produced across industries might show more fluctuation compared to the proportion of defective products in the quality area. This concept, akin to economists' consideration of market-related conditions, helps distinguish sectors that may be subject to different levels of volatility and fluctuation, hence reducing predictability of the future performances.

Despite these advantages, a notable drawback consists in the challenge of distinguishing whether the superior or inferior performance of KPIs in a specific sector is attributable to the consultants' capabilities or is intrinsic to the sector-specific characteristics. Therefore, a clear understanding of the type of intervention conducted by fischer Consulting's consultants is essential for evaluating sector-related performance and potential room for improvement in a specific project.

A critical complication arises due to the lack of any information regarding the methodology used for data collection, including the measurement techniques, instrument variability, and other influential factors. These aspects, discussed in depth in subsection 1.3 titled "Boundaries of data gathering process: origins, accuracy, and means of measurement," contribute to an unknown variance within our system, which is hardly predictable. This significantly reduces the ability to predict and comprehend specific performance outcomes, as well as understanding their underlying causes. The previous uncertainty is far from irrelevant since it has profound impacts on both the predictability and robustness of our model.

Now that we have characterized the methodologies and peculiarities of the dataset, we can proceed to clarify the specific categories considered and rationale behind their inclusion.

3.4.1. Types of categories

After an extensive series of research and consultations with the consultants of fischer Consulting, it was developed a comprehensive set of categories for identifying the KPIs under study, within a specific category, and it was structured in such a way that it will be suitable for potential future studies, recordings, and analyses.

The nine identified categories are as follows:

- Efficiency: focused on enhancing the utilization of existing resources, particularly by reducing the quantity of resources needed for a fixed amount of output.
- Productivity: associated with the volume of goods that can be produced within a specified time frame, irrespective of the resources employed.

- Quality: involves improving the quality of output by reducing defects produced by machines or humans in response to a set amount of goods produced or hours worked.
- Delivery: concerns the timely and accurate delivery of goods, extending to an examination of suppliers, whether internal or external.
- Costs: involves tracking and typically reducing production or function-related costs.
- Safety: encompasses tracking factors crucial for assessing safety levels, such as the number of injuries.
- Satisfaction: relates to the level of customer satisfaction and appreciation for current goods and services.
- Innovation: focuses on the company's ability to generate new solutions, including research and development of products and patent-related activities.
- Profitability: concentrates on increasing the overall profitability of the company.

These categories are designed based on macro-areas rather than highly specific ones, as sector specific KPIs. This approach proves particularly effective for fischer Consulting Italia, given the cross-sector services that they provide and the need to interface with clients from diverse background.

3.5. Tracked KPI

Throughout this study, was monitored a set of 15 Key Performance Indicators (KPIs) across four companies, accumulating a total of 1363 observations. It is noteworthy that these KPIs were exclusively associated with three crucial areas: productivity, efficiency, and quality.
In the following sections, we will delve into a detailed analysis of each distinct KPI, providing a breakdown for each individual customer.

It follows a summarizing table:

Firm's Name	KPI's Name	KPI description	KPI's Area
CLIENT1	KPI1	Losses during testing	QUALITY
CLIENT1	KPI2	PLT (Process Lead Time) projections	EFFICIENCY
CLIENT1	KPI3	Average Productivity	PRODUCTIVITY
CLIENT1	KPI4	Right First Time	QUALITY
CLIENT2	KPI5	Nutritional Samples Delay	EFFICIENCY
CLIENT2	KPI6	Bromatology Samples Delay	EFFICIENCY
CLIENT2	KPI7	TAT (Turnaround Time) Nutriline	EFFICIENCY
CLIENT2	KPI8	TAT (Turnaround Time) Bromatology	EFFICIENCY
CLIENT2	KPI9	Number of Fiber Samples Processed	PRODUCTIVITY
CLIENT2	KPI10	Number of Fiber Samples to be Redone	QUALITY
CLIENT2	KPI11	Number of Fiber Samples in Queue	EFFICIENCY
CLIENT3	KPI12	Average Mold Change Duration	EFFICIENCY
CLIENT4	KPI13	Production Delayed or On Time	PRODUCTIVITY
CLIENT4	KPI14	Rate of Production Delayed and On Time	EFFICIENCY
CLIENT4	KPI15	Backlog Value	PRODUCTIVITY

Table 2: KPIs description associated with each client and categorical distinction

3.5.1. Overview CLIENT n. 1

This client is a subsidiary operating under the control of an Italian company, a prominent player in the European heating and air-conditioning market.

Our analysis of this subsidiary encompasses four KPIs: losses during testing, Process Lead Time (PLT) projections, average productivity, and Right First Time (RFT).

Losses during testing represents the quantity of pieces lost during the setup testing phase across all machines at the site. While testing is crucial to ensuring correct machine setup before beginning the batch production, it impacts costs and raw material wastage, thereby afflicting the company's profitability.

Next, it was considered PLT, also known as Process Lead Time, which calculates the ratio between Work-in-Progress (WIP) and the exit rate (number of finished products produced in a given timeframe). This ratio, often referred to as Little's Law, is fundamental as it represents the time a raw material takes to traverse the entire process and become a finished product.

The third metric, average productivity, measures the number of sellable finished goods in relation to the total hours worked by all employees. This metric is significant as it accounts for rework, considering only those goods deemed marketable.

To address issues raised earlier, it was also examined the RFT, or Right First Time, which measures the percentage of finished products without defects relative to the total number of goods produced. This metric provides insight into the efficiency of the production process and the overall quality of the output.

3.5.2. Overview CLIENT n. 2

The second client holds a global presence in the market, specializing in sample analysis for various sectors, including food products, supplements, materials in contact with food (MOCA), cosmetics, and pharmaceutical products.
The first and second client's KPI are very similar. While the third is linked to percentage of samples, specifically Nutritional samples. The other KPI, it pertains to Bromatological samples.
The third and fourth KPIs are particularly relevant as they measure the Process Lead Time (PLT), also known as Turnaround Time (TAT). This metric signifies the time required to fulfill a request, providing crucial insights into operational efficiency.
The fifth and sixth indicators revolve around fiber samples. The fifth KPI reflects the number of fiber samples requiring rework, indicating areas for improvement in the number of quality sufficient analyses. The last indicator pertains to the number of fiber samples in the production queue, representing orders that have been registered but are awaiting production.

3.5.3. Overview CLIENT n. 3

The third client is a player in the manufacturing of plastic materials, specifically focusing on polyethylene grades PE100 and PE100-RC designed for water, gas pressure pipelines, and industrial fire-fighting applications.
Their KPIs centers around tracking the average changeover time of the stamp required to initiate the production of a new batch of products. This encompasses modifications in mechanical settings and potential computer adaptations for the subsequent batch. Notably, these measurements were taken both before and after implementing a Single Minute Exchange of Die (SMED) activity on the production machines.

3.5.4. Overview CLIENT n. 4

The fourth client stands as a pioneer in the realm of plastics and polymeric materials processing, evolving over time into a prominent figure in the core materials sector.

Concerning its first KPI, our focus was on tracking the number of panels, their final product, experiencing delays compared to the scheduled production. This aspect is particularly significant as fischer Consulting was engaged to minimize these delays and explore potential solutions for improvement in their management.

The second KPI, in contrast, monitored the ratio between delayed panels and those on schedule, relative to the planned production. This metric proves valuable, allowing to conduct precise and frequent observations to assess the current situation, accounting for seasonality and overall production dynamics.

The third and final metric pertains to the backlog value, representing the total value of orders yet to be fulfilled. Following a similar logic, mitigating the number of panels in the backlog could yield short-term revenue gains, thereby enhancing organizational efficiency.

3.6. The target achievement unit of measure

A critical aspect of our analysis implicated the formulation of the Target Achievement (TA) value for each recorded observation. Understanding the purpose behind its development and its utility is fundamental to comprehend the overall analysis.

The Target Achievement is defined as the percentage representation of the extent to which the desired target was attained at the moment of recording the related observation.

This definition hinges on two essential prerequisites: the existence of a target for each specific KPI and a reference value against which improvement or deterioration can be computed. Consequently, a target needed to be defined for each KPI, if not already established. This process involved consultations and discussions with consultants actively engaged in the project during the data collection, or sharing, phase. Notably, while the KPIs related to clients one and two had predefined targets set by the clients themselves, those related to clients three and four had targets internally defined by fischer Consulting members.

As for the second reference value, called *baseline* throughout the study, it was computed, when not explicitly stated, as the mean of all observations recorded before commencing collaboration with fischer Consulting. Although this metric may not always precisely reflect the central tendency of values, or being similar to the median value, it proved, in this study, statistically significant and indicative of overall performance before project initiation. This determination followed a detailed assessment of each KPI's specific variance and median values, also summarized by using box plots and compared to the respective mean.The cases where the baseline was not predetermined, as for clients three and four, the mean was adopted.

3.6.1. Target achievement computations and its advantages

The computation of target achievement was designed to yield a value of 100%, or 1 in decimal, when the desired target was fully met. Specifically, if the target exceeded the baseline, the target achievement would be the ratio between the differences of the observed value and the target in the numerator, and simultaneously, the target and baseline in the denominator:

$$TA = \frac{(observed\ value) - baseline}{target - baseline}$$

Conversely, if the baseline exceeded the set target, meaning that the aim to lower the metrics, the target achievement was computed as:

$$TA = \frac{baseline\ -\ (observed\ value)}{baseline - target}$$

It is evident from these formulas that when the observed value equals the baseline, the TA is valued at zero. It is important to note that while these metrics allow for unconstrained negative and positive values, in practice, fischer Consulting, like other consulting firms, imposes an upper cap, as is the case with client 1’s contract. Additionally, there is typically a lower limit of zero, especially since, most of the time, a variable fee, which is a portion of compensation whose final amount is dubious at time of agreement, because tied to certain future conditions, can be exchanged for a smaller, yet guaranteed, fixed fee or amount, whose final value not subject to any change since the time of agreement.

Additionally, there is typically a lower limit of zero, particularly because variable fees, constituting a portion of compensation billed to the client, kick in only when specific conditions are met to a certain extent. This sum can be substituted for a smaller but assured fixed fee or amount, the ultimate value of which remains unaffected by the passage of time or potential fluctuations in performance

The advantages of developing target achievement are numerous. Firstly, it unifies units of measure, disregarding variations or transformations used in computing each KPI. Secondly, this facilitates the comparison of different observations, even when measuring distinct elements in different units. Thirdly, it simplifies the comparison of achievement levels within a specified period across various KPIs. Lastly, it makes it easier to visualize the gap that needs closure to reach the desired target throughout the whole project. This is pivotal because the initial gap, as measured by TA, between the latest observation before the consultants' intervention began and the required target, may vary across different KPIs due to factors like higher volatility or the relative difficulty of achieving the target in relation to the additional compensation promises.

3.7. Evaluation of societal risk and risk averseness

The subsequent step in evaluating the feasibility of implementing a potential variable compensation involves examining the overall inclination of fischer Consulting, as well as its individual members, towards undertaking risky activities.

The significance of this assessment is underscored by the hypothetical scenario discussed in subsection 3.2.1, titled "Economical perspective on performance-related compensation package." In this hypothetical world, the consideration for compensation would solely be linked to the improvement in performance. However, as already discussed, theoretical concepts often collide with the imperfections of reality. Therefore, we must consider the possibility that fischer Consulting Italia may be unwilling to tie up all its compensation to the KPI level achieved. This reluctance is justified, as the company may lack influence over external factors, or the methodology used to measure and compute KPIs.

Considering this, it is reasonable to assume that fischer Consulting may only be open to exchanging a portion of its fixed income for a more uncertain variable compensation only when there is another upside. Therefore, the objective of assessing the risk aversion is to calculate the extent to which fischer Consulting members are willing to make the trade-off between certainty and uncertainty, along with the associated compensation that could be gained.

When evaluating a company's risk, there are multiple techniques to assess societal risk. One approach involves estimating the management's predisposition to take risks, while another involves conducting a thorough analysis of the company's financial records to evaluate its exposure to financial risk.

3.7.1. Assessing the risk through financial statement analysis

Assessing a company's financial risk exposure through the use of tools and ratios in financial statement analysis provides an objective means to evaluate societal risk. Despite its advantages, this approach has limitations, because of which a comprehensive explanation of the risk is not always possible. The following paragraphs will delve into the reasons behind this conclusion.

One primary challenge is data availability. For instance, publicly traded companies are bound by stringent disclosure requirements, safeguarding investor protection and market interests. However, such requirements are not required by law, with the same extent, for privately held companies, like fischer Consulting Italia. Unlike their public counterparts, private companies involve private parties in acquisitions or dispositions of ownership, hence capital at risk. Since no public interest is at stake, then there is no need for external regulatory bodies such as the SEC in the US or Consob in Italy. Therefore, private parties are able, within legal limits, to negotiate deals that align with their best interests.

In the case of fischer Consulting Italia, operating as a "Società in Accomodita Semplice" (Limited Liability Partnership), the lack of public disclosure requirements may restrict access to its financial statements to only a select few individuals. Consequently, these statements might not be readily available to analysts, impairing their ability to conduct in-depth analyses.

Another significant challenge is the lack of comparable entities. As a relatively young player in a niche market, fischer Consulting faces few competitors, none of whom have public disclosure requirements for financial statements. This absence of comparative data makes evaluating fischer Consulting's financial performance or risk exposure difficult, obstructing the computation of metrics like the Sharpe ratio or beta.

Considering that fischer Consulting was born only in 2017 and it has recorded an annual growth of approximately 30%, precision in assessing its current risk level should also account for the forecasted revenues, profit, and indebtedness level. Furthermore, the startup phase doesn't offer a representative historical data set, on which infer future trends.

3.7.2. Assessing the risk through individual risk averseness

A different approach to examining risk propensity for a company involves evaluating the inclination of both management and employees to embrace risks within the firm. Although this method is somewhat subjective compared to the previous one, its validity, relies in the idea that day-to-day decision-making relies in the hands of those in charge of operational activities rather than the owner himself. As a result, by tracing the risk-taking tendencies of fischer Consulting's team members, a broader understanding of the organization's overall risk disposition can be determined.

A common way to approach the evaluation of the propensity to take on risks, consist in submitting surveys whose questions have been already employed in other one. For instance, a survey could be design by drawing inspiration from those of the Socio-Economic Panel (SOEP), which is one of the largest and longest-running provider of multidisciplinary household surveys globally. Specifically, you could modeled your questions after those examined in DIW Berlin's paper 511, titled "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," authored by Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner, and published in September 2005. By adopting this approach, you could evaluate your results against a statistically significant and robust analysis of risk tendencies among individuals with varying characteristics. It's worth noting that the SOEP questionnaire used for analysis was administered to approximately 22,000 people in 2004.

3.7.3. Implementation of a risk-related model

In the search to establish a potential variable fee rate complementary to the fixed component, financial modeling emerges as a necessary tool to leverage. This modeling approach aims to factor in the adjustments required due to risk, which are inherent to both individuals and firms.

Central to the achievement of this goal can be the mean-variance model, which represents the base of modern portfolio theory, and that was developed by Harry Markowitz. This model presents a structured method for making informed investment decisions by balancing the two fundamental factors: expected return and risk.

Expected return denotes the anticipated gain or loss from an investment, or variable fee, within a defined timeframe. This is why implementing a solid understanding from a statistical point of view is necessary in guiding investors, or consultants, towards understanding the average outcome they can reasonably expect. Investors might make use of various factors such as historical data, economic forecasts, and market trends.

On the other hand, risk refers to the degree of variability of returns associated with an investment, or a particular project. Generally speaking, it represents the potential for a financial loss deriving from market fluctuations, economic downturns, or unforeseen events impacting specific assets, securities, or as in our case the KPIs performance.

The mean-variance model makes use of these components to construct what is known as the efficient frontier. This frontier delineates the optimal set of investment portfolios that offer the highest expected return for a given level of risk or the lowest risk for a given level of expected return.

In the evaluation of the introduction of variable compensation into the projects portfolio, the mean-variance model emerges as a valuable tool for evaluating associated risks. Here's how it can be utilized:

1. Risk assessment: the mean-variance model makes it easier to examine the risks inherent in various variable compensation options. By quantifying risk through metrics such as standard deviation and covariance, the model provides insights into the potential volatility and downside exposure of each particular project. This enables decision-makers, or the General Manager in fischer Consulting case, to choose the level of risk he/she is willing to accept in exchange for a variable, but potentially higher, compensation.
2. Diversification benefits: one of the key principles of the mean-variance model is the importance of diversification in reducing portfolio risk. By means of rigorous analysis, the model identifies the optimal combination of assets that minimizes overall projects portfolio risk while maximizing potential returns. By incorporating variable compensation, even alongside traditional asset classes, such as stocks and bonds, fischer Consulting can leverage diversification to reduce the additional risk introduced by variable components.
3. Stress testing: the mean-variance model allows for stress testing of the portfolio under various scenarios, including adverse conditions. By simulating potential market shocks, or as in our case: unlikely project performances, and assessing their impact on projects portfolio return,

the decision-makers can gain a comprehensive understanding of the risks associated with variable compensation.

In conclusion, the mean-variance model serves as a robust analytical structure for evaluating the potential risks associated with introducing variable compensation into a portfolio of projects. By conducting valid risk assessments, leveraging diversification tools, and stress testing, the decision-makers can make informed decisions that should maximize the outcome that derive from the desire for variable compensation with prudent risk management practices.

4. CHAPTER 4: Results of statistical analysis

This chapter delves deeply into the analysis of various KPIs, employing a diverse range of methods and statistical techniques to provide the most effective and comprehensive understanding. Many times, it will be made the distinction between analyses conducted at the level of individual KPIs and those carried out at a macro level, focusing solely on the broader category or area to ensure a complementary perspective.

4.1. General overview of popular statistical metrics

When working with a dataset, it's crucial to gain the most complete possible understanding of the data, including all subcategories of KPI and the overall observations recorded. For this reason, initially, it was examined whether any variables were containing any missing values or if the dataset was complete. This was verified by simply counting the *non-NA* values for each column, where *NA* represents "not available" within the R environment.

The initial variables, derived from the dataset, were the following: *firm_name*, *prj_code*, *date_start_prj*, *date_end_prj*, *KPI_id*, *index_KPI*, *date_obs*, *performance*, *target*, *KPI_des*, *KPI_area*, *baseline*, and *target_achievement*. For a detailed description of each variable, please refer to the appendix, reference 1.

4.1.1. Balanced Data Analysis

While looking at any data which is related to an event study, it is important to assess whether the dataset is balanced in terms of observations before and after the project initiation, in this particular case. To accomplish this, a new variable, after, was created to determine if a given observation was recorded before or after the project start date (*date_start_prj*). Combining this variable with all simple observations count yielded the following output:

	CLIENT 1	CLIENT 2	CLIENT 3	CLIENT 4
After	161	433	38	22
Before	116	260	320	13

Table 3: Number of observations for each client, before and after the commencing of the project

The table clearly indicates a significant imbalance in the number of reported observations before and after the project initiation for client 3, especially concerning the KPI: "Average Mold Change Duration." Given the sheer volume of data, it's strategic to retain this KPI, because of its historical significance. This means that if any model were to be implemented such that it relies on historical data, then this KPI can offer valuable insights into the relationship between expected and actual values. Although this won't directly influence the techniques used for data analysis and modeling, it afflicts our analysis by allowing us to discern more accurately the potential causes of unsolidity in our results. This is particularly crucial for extensive analyses such as the time series, where it becomes important the sheer of data upon which the model can be computed.

KPI's Name	Count
KPI1	39
KPI2	104
KPI3	54
KPI4	80
KPI5	39
KPI6	39
KPI7	39
KPI8	39
KPI9	179
KPI10	179
KPI11	179
KPI12	358
KPI13	12
KPI14	12
KPI15	11

Table 4: Number of observations for each KPI

Another note of caution is justified when evaluating the data for client 4, which has three different KPIs and a total of 35 observations. This relatively small sample size may challenge the statistical significance of these observations throughout the analysis.

In the other cases, while imbalances exist, they are less concerning, provided that sufficient data is available for each KPI. However, as evident in Table 1, the data identified as KPI13, KPI14, and KPI15 of client 4 may prove insufficient.

Moreover, to gain a wider understanding of data balance, it's essential to examine the balance of categories. The following displays the number of observations for each category, before and after:

KPI's Area	After 1 = Yes 0 = No	Count
EFFICIENCY	0	625
EFFICIENCY	1	351
PRODUCTIVITY	0	113
PRODUCTIVITY	1	155
QUALITY	0	22
QUALITY	1	97

Table 5: Number of observations for each category present, before and after the commencing of the project

Although there is a noticeable imbalance between before and after observations for *EFFICIENCY* and *QUALITY*, the available data should still allow us to draw conclusions with a reasonable degree of certainty.

Hence, while on one hand, it's important to acknowledge imbalances that may impact the robustness of our analysis and the validity of the results, in this case, it is possible to proceed with our analysis, taking into account those highlighted imbalances. Nonetheless, it's prudent to interpret the results with some reservations.

From a managerial perspective, it would be prudent to conduct analysis across a wide array of KPIs that are directly impacted by the consultants' intervention. This approach is driven by two key reasons:

4. Robustness of results: increasing the number of data points available for analysis reduces the likelihood of encountering non-representative KPIs. Moreover, as the dataset expands, the weight that this biased KPIs exercise on the overall findings results within a given area falls.
5. Negligible variable costs of software: while developing software can be time-consuming and expensive, when executed effectively, it can be utilized regardless of the number of KPIs it needs to analyze. Consequently, there are no additional costs associated with analyzing an additional KPI. For example, the software developed for this thesis could handle 100 or 15 KPIs with equal efficiency and effort. This capability is due to the automated saving of analyses, graphs, and results in specific type in predefined folders, enabling seamless adaptation or filtering of data as required.

4.1.2. Descriptive statistics for performance and target achievement

When analyzing a substantial dataset, it's essential to compute key descriptive statistics to gain insights into the different variables' characteristics, in this case of performance and target achievement. Three basic metrics to calculate are the mean, median, and standard deviation, because they offer valuable intuitions into the central tendency and the dispersion of the data.

In this first part, a bottom-up approach was used, by starting with individual KPIs and differentiating them into two phases: pre-project and post-project. In reviewing the resulting table (Appendix, Attachment 2), it's evident that some specific KPIs exhibit relevant complexities. Notably, KPI2 and KPI10 showed significant variability compared to their respective means and medians. While KPI2's variability is concerning (above 50% of the mean) it remains within reasonable bounds. On the other hand, KPI10 displays a worrying variability exceeding 150% of its mean value. This anomaly is multiplied when considering the TA values, which show wildly variability, especially on the 28th of February, where the TA reached a value of -7300%, as illustrated in the following graph.

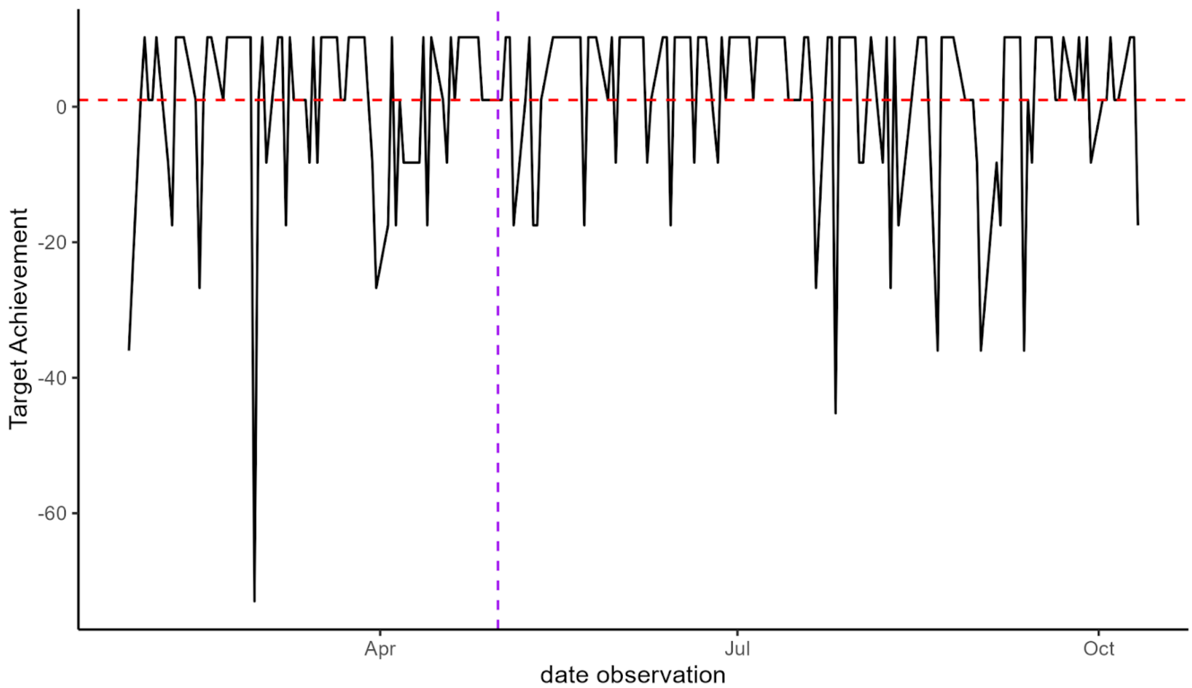

Figure 8: Line graph of Target Achievement for date of KPI10

The inexplicable variability within KPI10 can be attributed to its measurement criteria, baseline, and target. With a baseline of 1,108 and a target of 1, KPI10, which tracks the "Number of Fiber Samples to be redone," had values ranging from 0 to 9. This wide range meant that a single additional unit of sample that required rework could lead to a substantial variation in TA. Moreover, while the maximum TA was capped at 1026%, there was no lower bound restriction, allowing TA to plummet to extremely low values. Here is a graph of the number of sampled that needed rework:

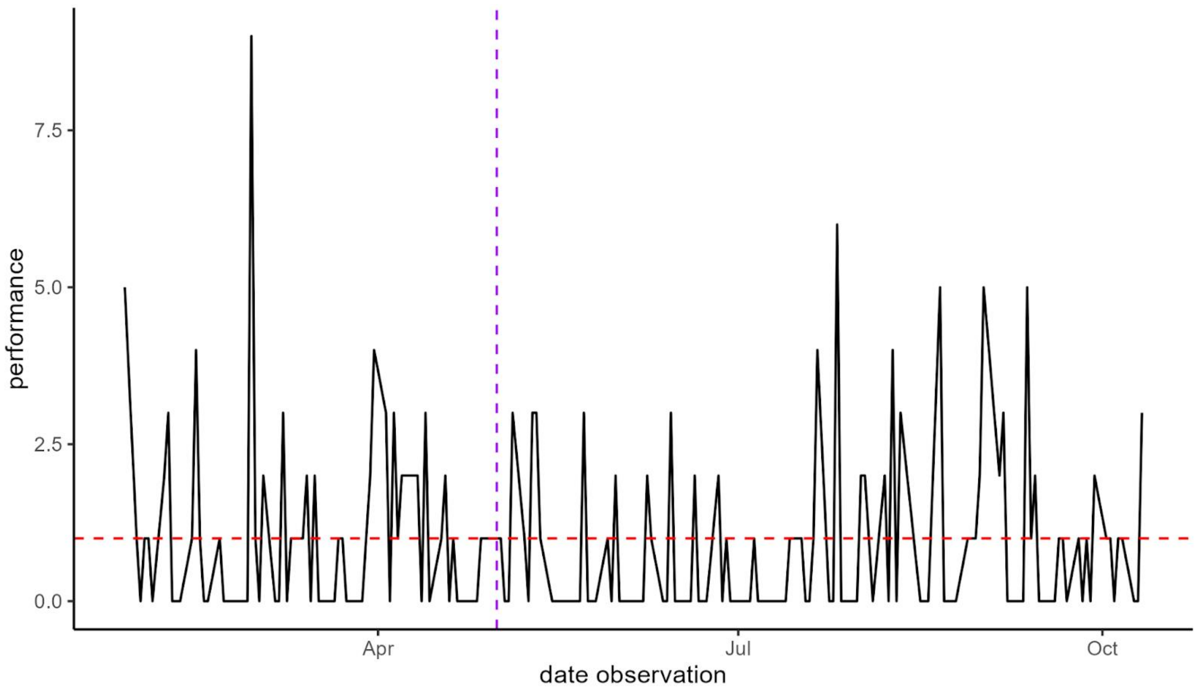

Figure 9: Line graph of Performance for date of KPI10

Given these findings, it was decided to exclude KPI10 from further statistical analysis due to its disproportionate influence on TA. On the other hand, despite the variability observed in KPI2, its consistency suggests the following of a discernible trend rather than a random fluctuation, which calls for greater attention and prevents its premature exclusion from the dataset. Subsequently the graph of performances for KPI2:

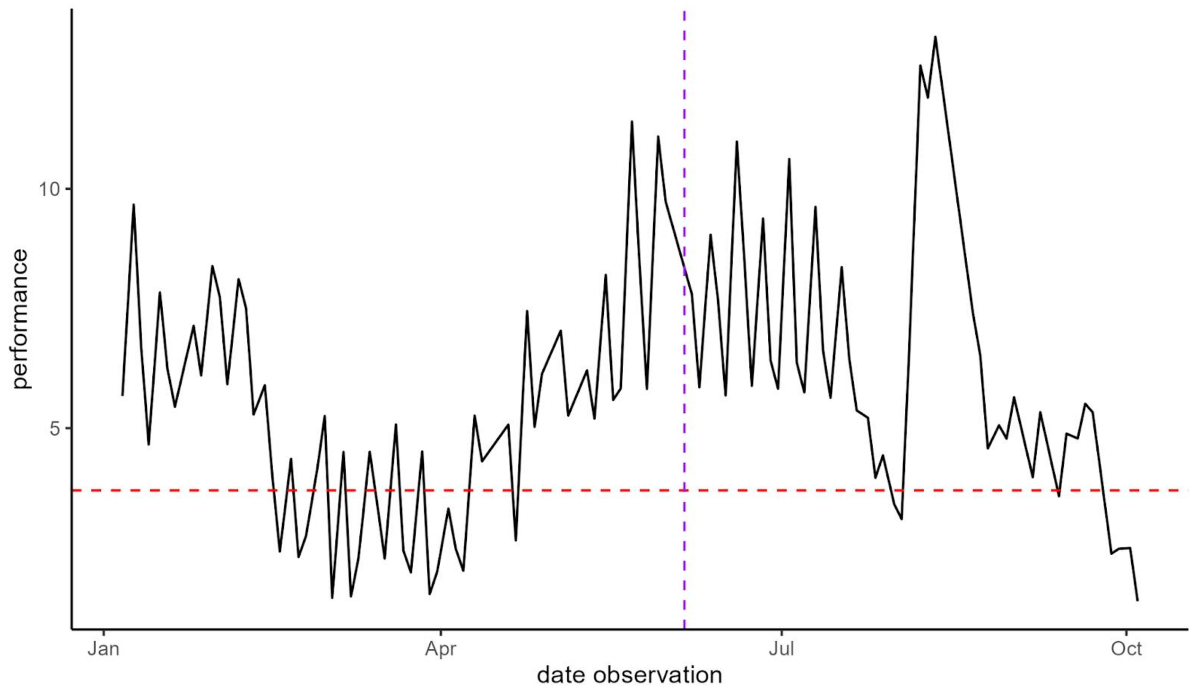

Figure 10: Line graph of Performance for date of KPI2.

Expanding our analysis, we utilize TA to compare achievement levels across projects and areas. Initially, an equal-weighted average was computed, allowing for all projects to have the same level of influence. Subsequently, a weighted average was calculated, assigning weights based on the proportion of observations recorded. Both averages were distinguished based on whether they were calculated before or after the project commenced.

Equally weighted averages:

	Mean Target Achievement	Std. Deviation of TA
After	57,99726%	76,92523%
Before	- 10,7159%	67,98682%

Table 6: Equally weighted averages of Target Achievement before and after the commencing of the project

Observations weighted averages:

	Mean Target Achievement	Median Target Achievement	Std. Deviation of TA
After	34,51601%	41,85437%	142,6634%
Before	0,336910%	6,819650%	152.0265%

Table 7: Observations-count weighted averages of Target Achievement before and after the commencing of the project

The weighted averages demonstrate a smoothing effect on the TA differences before and after the project with respect to the equally weighted one, while still maintaining a positive difference. However, one intriguing shift is in the mean of TA, which moves from negative to positive values for those observations recorded before. This shift prompted reflection on the motivation behind it, since it will make little sense to engage a consulting firm when desired output has been already attained, unless the scope is to reduce the variability within it.

Of particular interest is the examination of standard deviation. While initially alarming, the level of standard deviation in this and subsequent analyses highlights the challenges in achieving precise predictability of individual observations. Therefore, while a general trend may be discerned and computed, it is crucial to employ cautious interpretation and further refinement of the data, as to improve individual predictive capabilities.

As a manager, I would find these partial results noteworthy, as they highlight a considerable shift in the average level of achievement since the project began. The improvement is significant, nearly 70% across all KPIs and 34% when weighted. Although still distant from the desired outcome of 100%, this progress is a promising starting point, which underline the positive impact that fischer Consulting has had on KPI performances. However, the notably high level of standard deviation call for caution and further investigation, so any prudent manager should beware before committing time and resources to promising returns of 70% across all interventions on any KPI to any client.

To explore instead the potential differences between the mean and median values, which seem to suggest an asymmetry in the distribution of TA, a categorized value of TA was computed, dividing the scores into bins of 50%, with a range of values of $\pm 350\%$. The resulting insights are visualized in the following bar graph:

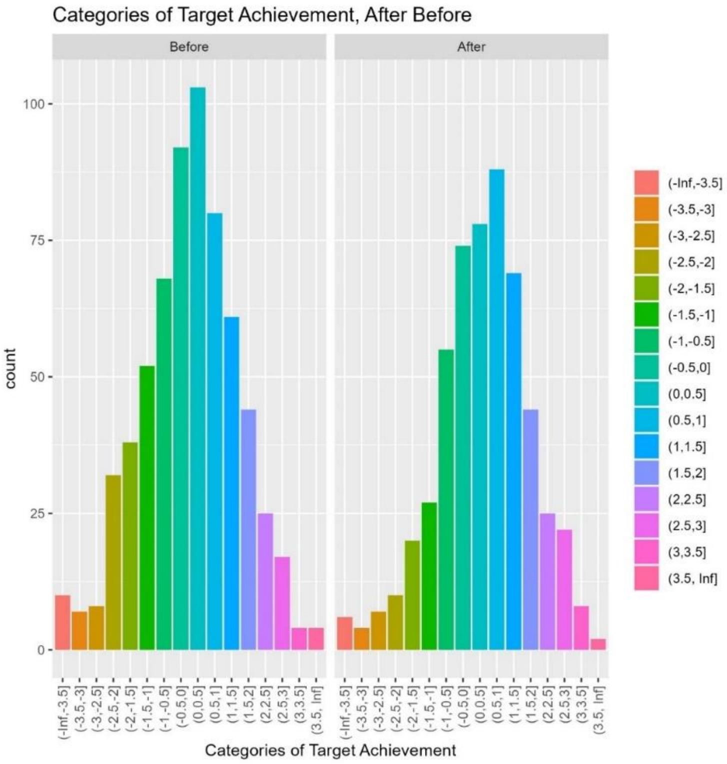

Figure 11: Bar graph of categorized Target Achievement before and after the commencing of the project

The graph confirms a slight rightward translation of the central tendency in the distribution for values observed after. Nevertheless, it's evident that the distribution is somewhat skewed to the right in the left panel, potentially explaining the discrepancy between mean and median values. A box plot was omitted due to its limited ability to elucidate data distribution peculiarity and its challenge in discerning differences between before and after observations.

Upon the exclusion of KPI10, the distribution of observations across different areas has changed, as a total of 179 observations were removed from further analysis. The new balance by areas is as follows:

KPI's Area	After 1 = Yes 0 = No	Count
EFFICIENCY	0	510
EFFICIENCY	1	287
PRODUCTIVITY	0	113
PRODUCTIVITY	1	155
QUALITY	0	22
QUALITY	1	97

Table 8: Number of observations for each category present, before and after the commencing of the project, without KPI10

It's crucial to note that KPI10 was initially part of the efficiency area.

4.1.3. Descriptive statistics for categories

While addressing the individual trends was a crucial task, it is equally important to assess them within a broader context which is defined by the categories. Similar to the previous analysis of the KPIs, a thoughtful evaluation of the overall statistics and distribution of target achievement was conducted for each area.

The initial phase involved assessing the mean, median, and standard deviation of these areas, before and after the consultancy took effect. The results table are presented here:

KPI Area	After 1 = Yes 0 = No	TARGET ACHIEVEMENT			
		Mean	Median	Std. Deviation	n. Obs.
EFFICIENCY	0	0,02	0,13	1,50	510
EFFICIENCY	1	0,51	0,71	1,70	287
PRODUCTIVITY	0	-0,08	0,04	1,74	113
PRODUCTIVITY	1	-0,14	-0,08	1,01	155
QUALITY	0	0,00	0,09	0,54	22
QUALITY	1	0,61	0,63	0,79	97

Table 9: Mean, median, std. Deviation, and n. of observations for each KPI area before and after commencing the project

Overall, the results exhibit two distinct trends: an improvement in overall performance coupled with higher variability, or a decline in performance correlated with reduced variability.

Of the former scenarios are part efficiency and quality. They demonstrate a marked improvement, with their means moving from an initial near zero value to surpassing 50% of TA. This suggests a tangible positive impact for the consultancy advice. Similarly, the median follows a close trajectory to the mean. However, a notable increase in standard deviation was observed. The efficiency category recorded was above 150% of the settled target, which raises questions about whether this shift is due to chance or a causal effect, even though it must be noted that while the variance observed an increase of 13%, the mean observed change of 2450%, a significant difference. Moreover, the abundance of data makes it statistically improbable to be solely driven by randomness.

On the other hand, while quality exhibits a higher mean increase than efficiency, the statistical significance of this result is debatable due to the limited number of observations, particularly in the before period. Nonetheless, the substantial change seems to suggest that the impact of the consultants' interventions was indeed positive.

Of greater concern is productivity, where performance has decreased both in mean and median values. Significantly, alarming is the median's shift from positive before to negative after the project started. This should be the case in which a comprehensive examination and consideration of external factors is urgently needed, in order to fully grasp the underpinning causes of this change. Nevertheless, despite the negative impact on mean and median, variance has decreased by 42%, suggesting an honest reduction of variance, when considering the volume and balance of observations on both ends.

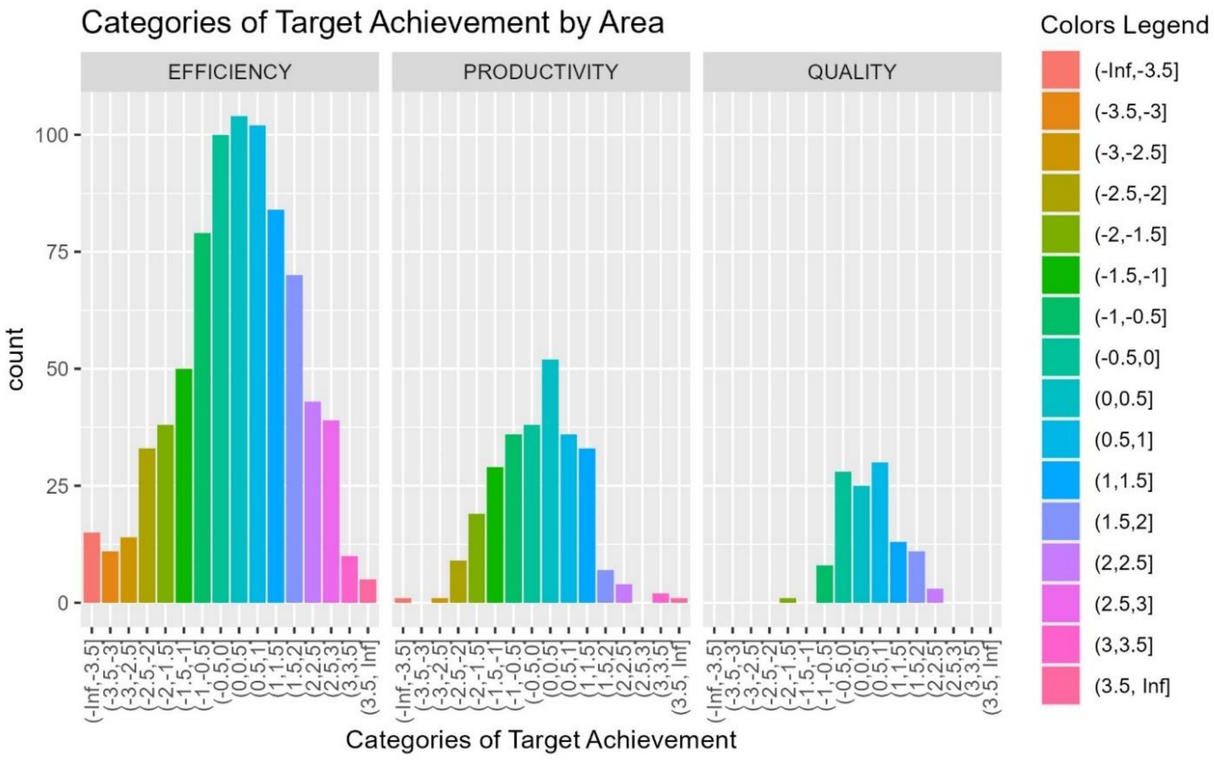

Figure 12: Bar graph of categorized Target Achievement before and after the commencing of the project by KPI area

After having received confirmation of these results through further analysis, a skilled manager should leverage this knowledge during negotiations and to address the existing performance gaps. Primarily, it's evident that significantly higher performances are obtained for KPIs related to either efficiency or quality. However, due to the substantial variability observed, a strategic approach for a manager would be to avoid focusing solely on specific checkpoints when determining variable compensation. Instead, using mean values would prove more effective, reducing exposure to individual performance volatility.

Moreover, it's essential to delve into the root causes behind the deteriorating performances in the productivity area. This investigation should clarify whether these issues are due to the unique aspects of the project or if they reflect broader inefficiencies in the approach to production enhancement consultancy. Such insights are invaluable as they guide managers and teams towards focusing their efforts on achieving the desired outcomes effectively.

4.2. Standardization and Z-Score

4.2.1. General overview of Z-Score

In the initial phase of analyzing the Z-Score, calculated for each individual observation relative to its corresponding KPI, the emphasis was placed on assessing its general trend and characteristics. Specifically, there was a decision not to evaluate the Z-Score of the TA, as doing so would introduce redundancy. Since the Z-Score provides a standardized measure that can be compared across different KPIs, considering also the TA would duplicate information. Moreover, since the TA is derived from performance metrics, its Z-Score would mirror that of the performance Z-Score, offering no additional utility and thus causing redundancy.

The primary assessment involved analyzing the distribution of Z-Score values to gain insights into their variance and pattern. Below is a summary bar graph illustrating this distribution:

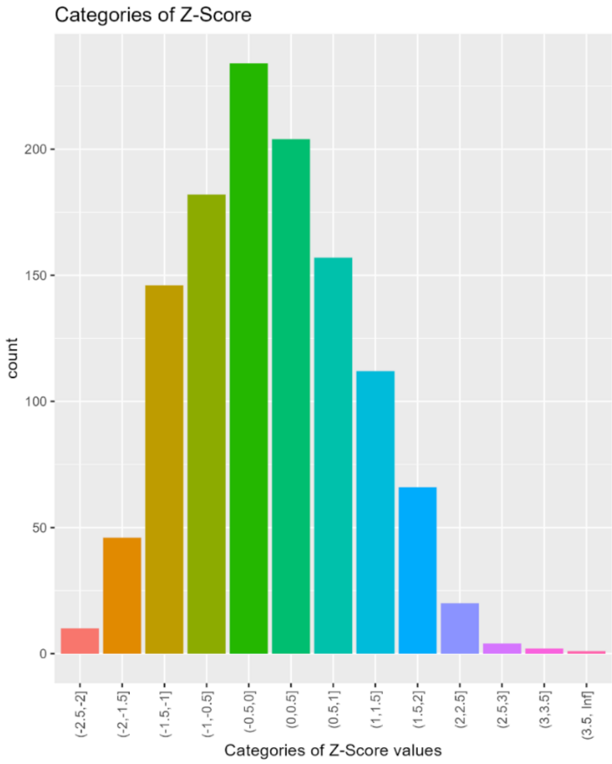

Figure 13: Bar graph of categorized Z-Score of all observations

Z-Score			
Mean	**Median**	**Std. Deviation**	**n. Obs.**
-1.02E-16	-0.07835161	0.98894990	1184

Table 10: Mean, median, std. Deviation, and n. of observations for the Z-Score of all observations

As underlined from the data and confirmed by the median values, the distribution of Z-Scores shows a left-skewed pattern. This suggests that if this distribution persists across each KPI, we can anticipate a higher concentration of values on the left side of the mean, which is zero, but also a lower range of distribution of these values with respect to those on the right side of the mean. Hence, it's not going to be surprising if we will find a negative median for each KPI falling within the range of -0.5 to 0.

It is important to note that the mean is approximately zero, and the variance is nearly one, in both cases. These characteristics are a consequence of the computational structure employed to derive the Z-Score, and if computed correctly we shall always find these.

By going deeper into the analysis by differentiating observations before and after the initiation of each project, we can observe a distribution of categorical Z-Scores as follows:

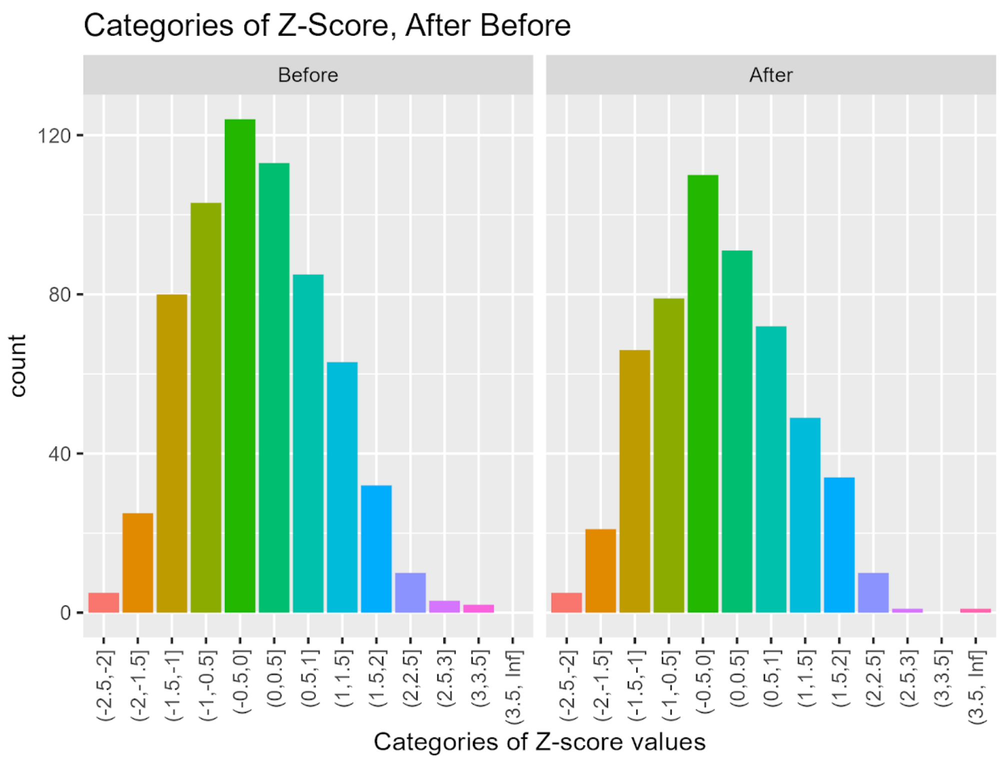

Figure 14: Bar graph of categorized Z-Score before and after the commencing of the project of all observations

After 1 = Yes 0 = No	Z-Score			
	Mean	**Median**	**Std. Deviation**	**n. Obs.**
0	6E-17	-0.0741756	0.9906394	645
1	-3E-16	-0.1001695	0.9878443	539

Table 11: Mean, median, std. Deviation, and n. of observations for the Z-Score before and after the commencing of the project of all the observations

By looking at the graph and the table, the previously discussed assumptions of the Z model hold true even in this scenario. Nevertheless, a significant change has occurred: the median has shifted from -0.10 before to -0.07 after. This shift indicates that after the project the standardized values are usually more evenly distributed along the sides of the mean, with respect to before. Consequently, this improvement in the symmetry of observed standardized values can improve the predictability of the problem, and by consequence of the performances.

4.2.2. Z-Score by KPI

This subsection is going to analyze the results of the Z-Scores analysis across the KPIs. The table below summarizes the median Z-Scores after and before the start of the project for each individual KPI, along with the number of observations recorded.

KPI id	After 1 = Yes 0 = No	Z-Score	
		Median	n. Obs.
KPI1	0	-0,17	22
KPI1	1	0,06	17
KPI2	0	0,00	58
KPI2	1	-0,20	46
KPI3	0	0,01	36
KPI3	1	-0,10	18
KPI4	1	0,05	80
KPI5	0	0,40	17
KPI5	1	0,08	22
KPI6	0	0,08	17
KPI6	1	-0,06	21
KPI7	0	0,08	17
KPI7	1	-0,04	23
KPI8	0	-0,13	17
KPI8	1	0,13	22
KPI9	0	-0,04	64
KPI9	1	-0,12	115
KPI11	0	-0,30	64
KPI11	1	-0,22	115
KPI12	0	-0,07	320
KPI12	1	0,00	38
KPI13	0	-0,36	5
KPI13	1	0,06	7
KPI14	0	-0,43	5
KPI14	1	-0,31	7
KPI15	0	0,22	3
KPI15	1	-0,07	8

Table 12: Median and n. of observations for the Z-Score before and after the commencing of the project by KPI id

While the table offers insights into the median Z-Scores, it doesn't explain the underlying distribution of the data. To gain that insight, it was computed the variation of each KPI's Z-Score through graphical representations. By examining these graphs, we can gain some additional information about whether the performance changes are statistically significant.

Ideally, in a normal distribution, we would expect the median to shift towards the mean, which is zero. This would indicate an improvement in performance predictability. Based on this assumption we categorize the results into three groups:

- Positive Improvement: KPIs where the initial gap between the median and zero has narrowed. These are KPIs 1, 5, 11, 12, and 13.
- Negative Improvement: KPIs where the gap has widened, indicating a deterioration in performance. KPIs 2 and 9 fall into this category.
- Impact to be Assessed: KPIs requiring further analysis due to unclear trends. Those remaining one: KPIs 3, 6, 7, 8, 14.

Let's take KPI1. The divergence bar graph, which is reported below, doesn't show a significant change between the pre-project and post-project periods. This observation holds true even for all other KPIs in the positive group, except KPI5.

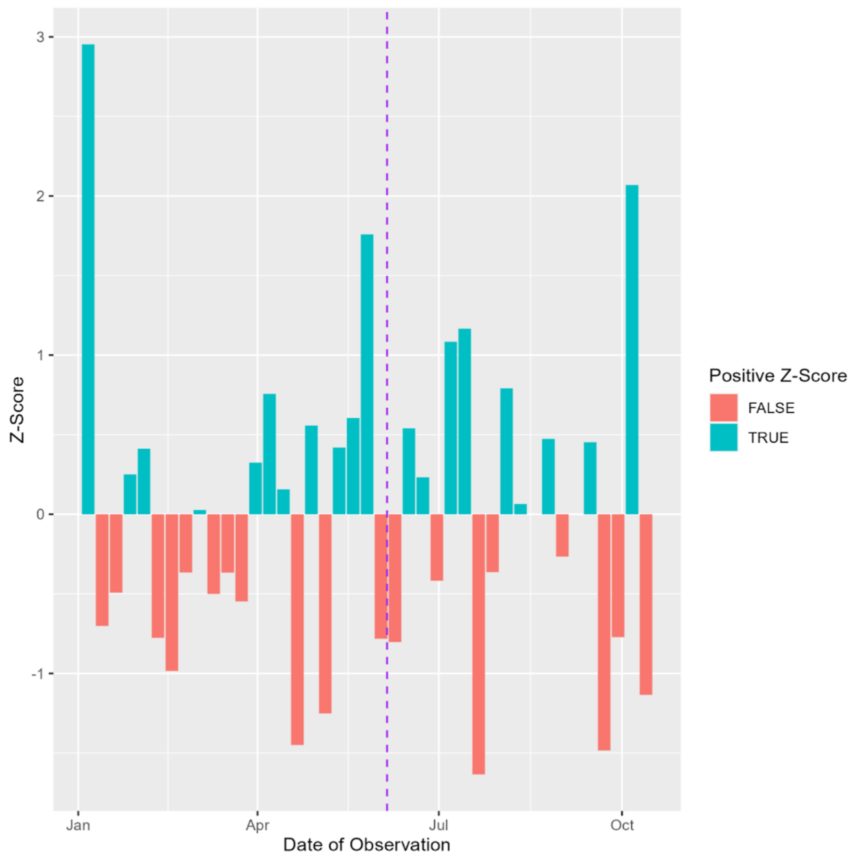

Figure 15: Divergence bar graph of Z-Scores for KPI1

For KPI5, here below, there's a noteworthy shift in performance. Initially starting with a high positive Z-Score, it progresses towards values closer to zero in the post-project timeframe, particularly on the positive sides. However, there's a random variance on the negative side, indicating inconsistency in sample delays. This might suggest a potential upper limit on sample delays, but no physical lower limit, which causes these unpredictable patterns.

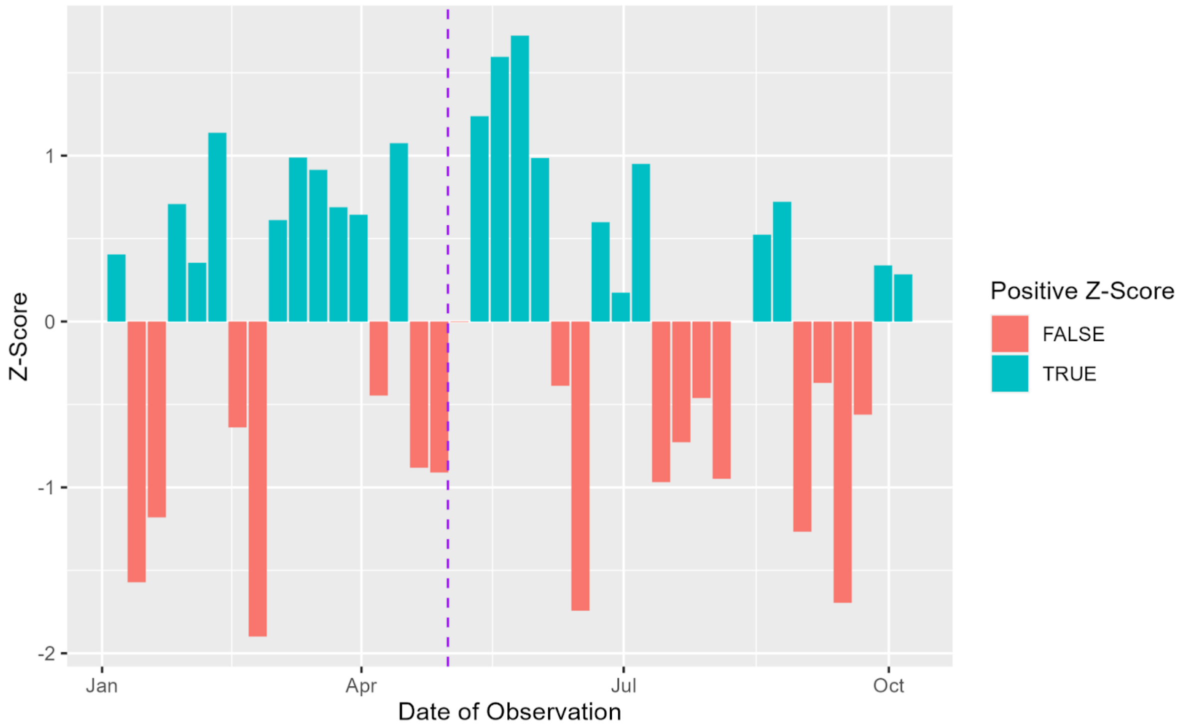

Figure 16: Divergence bar graph of Z-Scores for KPI5

Similarly on the negative group, doubts persist regarding the relevance of the median for KPI9, as illustrated by its graphical representation. More analysis should be needed to understand the underlying trends and potential factors influencing its performance.

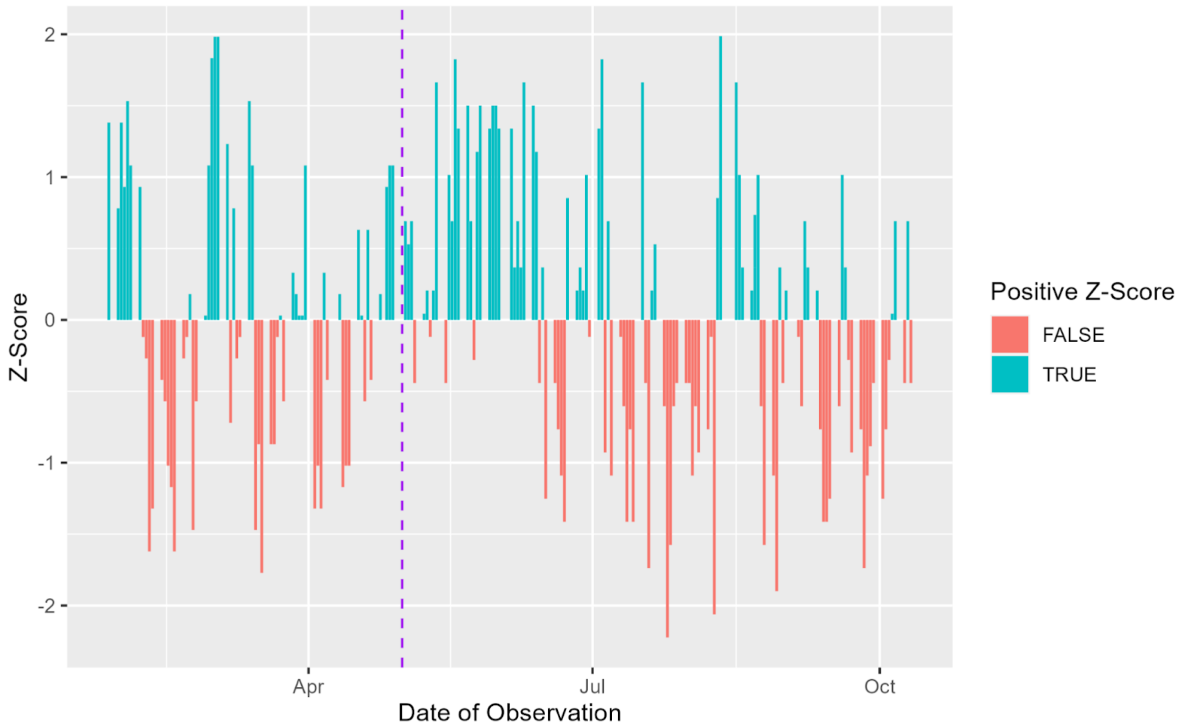

Figure 17: Divergence bar graph of Z-Scores for KPI9

The evaluation of KPI2, below, instead reveals a concerning negative trend, identified by two distinct plunges in the after period. Interestingly, the graph seems to indicate that Z-Score values change suddenly between assuming exclusively in the negative values to positive ones, while keeping this following positive, or negative, trend for some times, before another abrupt interruption shifts it to the opposite side of the zero. This behavior makes the median potentially insignificant for our research.

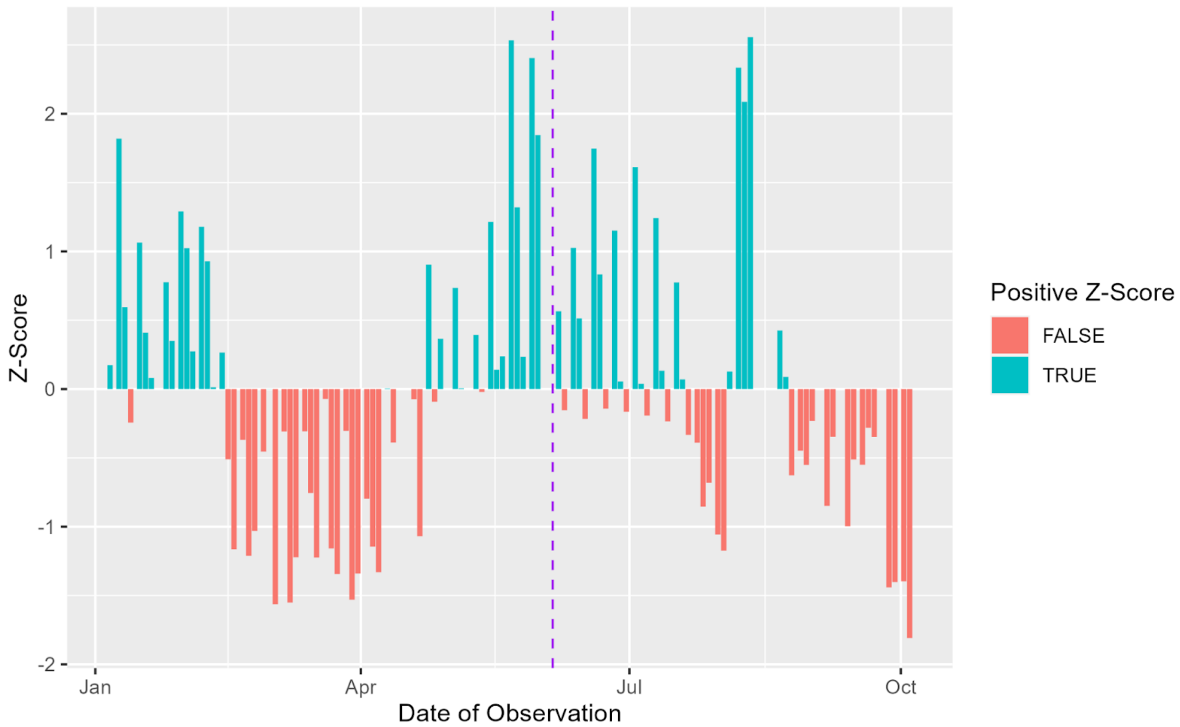

Figure 18: Divergence bar graph of Z-Scores for KPI2

The other KPIs which are observed have different particular situations. In the case of KPI6, KPI7, and KPI8, these were simply highly balanced, and negative values of Z-Score were balanced by similar positive ones.

KPI14 and KPI15, unfortunately, do not have enough observations which makes them not so relevant in the following context. KPI3 has observations which are inserted periodically, hence by lot, which cause it to have positive and negative values on the same day, which makes the assessment of the Z-Score still useful for historical observation and explanation, but further from being optimal for seeing and tracking progress throughout time.

For what concern KPI4 instead since it only contains values which are recorded after the commencing of the project there is no historical data against which we can compare those observations, hence the median has no useful application here.

As a manager, it's crucial to note that despite minimal observable change in the overall Z-Score from the project's onset to its conclusion, still computing the Z-Score can offer valuable insights. This metric provides a sense of the team's ability to normalize performance trends. While this normalization may not be the primary objective of the consultancy, it is still useful because, by normalizing the performances we can improve the accuracy of production planning, hence predicts costs associated with defects and overtime—both influenced by machine or human performance—and reduce the variability in performance. Moreover, more predictable performance freed up time for employees and employers to engage in other activities, because it eliminates the need to handle unforeseen circumstances. Leveraging this tool can significantly bolster fischer Consulting's market position and yield additional benefits.

4.3. Correlation

Part of the analysis comprehends the examination of the correlation between two key variables: *performance* and *diff_since_begin*, which denotes the time difference in days between the observation date and the project's start date. This measure includes both negative values, indicating observations before the project's commencement, and positive values, reflecting observations afterward.

The correlation analysis employed the *cor.test( )* function, which, as seen, is based upon Pearson's product-moment correlation coefficient to assess the relationship between the two variables.

Before proceeding, it is crucial to consider one already mentioned characteristic of our KPIs, which is relevant to our analysis. When evaluating performance, it's basilar to distinguish between those KPIs who have a target which is higher than their respective baseline. This comparison is identified by the *Target Higher* column, where *NO* means that the target is not higher, nor equal, to the baseline, and *YES* indicates that the target exceeds the baseline.

Firm's Name	KPI's Name	KPI description	Target Higher
CLIENT1	KPI1	Losses during testing	NO
CLIENT1	KPI2	PLT (Process Lead Time) projections	NO
CLIENT1	KPI3	Average Productivity	YES
CLIENT1	KPI4	Right First Time	YES
CLIENT2	KPI5	Nutritional Samples Delay	NO
CLIENT2	KPI6	Bromatology Samples Delay	NO
CLIENT2	KPI7	TAT (Turnaround Time) Nutriline	NO
CLIENT2	KPI8	TAT (Turnaround Time) Bromatology	NO
CLIENT2	KPI9	Number of Fiber Samples Processed	YES
CLIENT2	KPI11	Number of Fiber Samples in Queue	NO
CLIENT3	KPI12	Average Mold Change Duration	NO
CLIENT4	KPI13	Production Delayed or On Time	NO
CLIENT4	KPI14	Rate of Production Delayed and On Time	NO
CLIENT4	KPI15	Backlog Value	NO

Table 13: Target higher variable for all the KPIs and ancillary informations

It was anticipated that the observations of a positive correlation in KPIs where the target exceeds the baseline would've implied that as the *diff_since_begin* metric increased, we expected the *performance* to improve. Conversely, for KPIs where the target falls below the baseline, we anticipate a negative correlation, indicating a decrease in *performance* as the *diff_since_begin* metric increases.

An initial issue has arisen regarding the KPI analysis for client 4. Insufficient observations were available to meet the necessary constraints for computing Pearson's product-moment correlation through the use of the *cor.test( )* function. Consequently, our analysis is limited to the KPIs of clients 1 and 2.

After having looked at the correlations table, we discern two distinct scenarios, one for client 1 and another for client 2. Firstly, in the case of client 1, the p-values for KPIs 1, 2, and 3 surpass the 30% threshold, indicating statistical insignificance. However, KPI 4 deviates from this trend, and the *cor.test( )* yields the following results:

Pearson's product-moment correlation						
KPI id	**t statistic**	**degrees of freedom**	**p-value**	**95% interval**		**correlation**
KPI1	0,82779	37	0,4131	-0,1887	0,4132	0,1348
KPI2	0,97563	102	0,3316	-0,0983	0,2834	0,0962
KPI3	0,07701	52	0,9389	-0,2578	0,2776	0,0107
KPI4	-3,3866	78	0,0011	-0,5356	-0,1501	-0,3580
KPI5	-4,4795	37	6,95E-05	-0,7653	-0,3413	-0,5923
KPI6	-4,4761	36	7,35E-05	-0,7704	-0,344	0,5976
KPI7	-3,4463	38	0,0014	-0,694	-0,2081	-0,4880
KPI8	-2,1431	37	0,03875	-0,5863	-0,0188	-0,3323
KPI9	-4,1095	177	6,06E-05	-0,4235	-0,1552	-0,2951
KPI11	-1,4409	356	0,1505	-0,1784	0,0277	-0,0761

Table 14: results of the correlation analysis, Pearson's product-moment correlation

While the p-value remains a key consideration, it's also crucial to keep in mind the confidence interval, here computed at 95% significance. Analyzing this, we note that although, for instance, KPI8 remains statistically significant (p-value below 4%), its interval spans from nearly -0,60 to almost zero. This suggests that the correlation may be very close to zero, which mean that there might be no correlation at all, yet still yield a statistically significant result.

The last evaluation concerns the KPI11 of client 3, which shows a p-value which is insignificant, then we cannot reject the null hypothesis, hence we cannot consider the computed correlation, which in any case was already close to zero, to be statistically significant.

Additionally, it's worth noting the absence of certain KPIs, specifically KPI12, 13, and 14, due to insufficient observations to compute the correlation by making use of the *cor.test( )*. Consequently, we lack the means to assess the correlation between date differences and performance for these KPIs.

4.4. Linear Trend

One of the most useful analyses to conduct is the regression analysis, which emphasizes the correlation, and not, the causation between variables.

During various phases of this analysis, each KPI and area of performance was examined, employing a linear regression, explored in the previous chapters. This approach aimed to discover whether significant impacts could be observed and evaluated over time, by comparing the performance against the date of observations of these lasts.

4.4.1. Linear tendency by KPI and period of recording

In this subsection, other than evaluating the correlation between performance and the date of observation, it was distinguished between values recorded before the project's commencement and those recorded afterward.

For each individual KPI, a linear regression was computed using the *lm()* function in R, accompanied by a graphical representation illustrating the general trend.

Prior to the project's initiation, it can be anticipated that a near-zero linear correlation between performance and dates should be observed. This would signify that performance remains largely unaffected by external factors on a daily or even weekly basis. Thus, while it is expected that performances would fluctuate around a certain mean, with varying degrees of variability, they would linearly tend to it. However, we should not be surprised if the KPIs were to exhibit positive or negative trends, because, in the former case, the necessity of a consultant intervention might be even higher than usual, while in the latter while it might exist a positive trend this might not satisfy the necessary productive expectations, or its pace of improvement is unsatisfactory.

Conversely, following consultant interventions, it should be expected that performances show an improvement. Nevertheless, a complication arises as certain KPIs have different target benchmarks, some of which are higher than their baseline value and others which are lower. Consequently, we should not expect all coefficients of date of observation to trend uniformly upward or downward; instead, we should seek for concordance between coefficient signs and the relationship between baseline and target benchmarks.

Commencing with pre-project values, the following statistics are observed:

KPI id	Intercept	Coeff. Date Observation	p-value	R2	R2 Adjust	AIC	BIC	RMSE
KPI1	20.3114	-0.0008	0.8449	0.0020	-0.0479	60.6112	63.8843	0.8371
KPI2	-116.0410	0.0062	0.4106	0.0121	-0.0055	271.7912	277.9725	2.3926
KPI3	-1.6150	0.0001	0.8531	0.0010	-0.0284	-14.3474	-9.5968	0.1824
KPI5	-3.0882	0.0002	0.7086	0.0095	-0.0565	-41.4362	-38.9366	0.0600
KPI6	14.9746	-0.0008	0.1131	0.1569	0.1007	-41.1305	-38.6309	0.0605
KPI7	-41.7528	0.0025	0.4128	0.0449	-0.0188	22.1243	24.6239	0.3888
KPI8	31.1522	-0.0013	0.8156	0.0037	-0.0627	44.1146	46.6142	0.7424
KPI9	499.8641	-0.0245	0.4427	0.0095	-0.0064	428.7055	435.1822	6.5761
KPI11	-296.5973	0.0230	0.8801	0.0004	-0.0158	629.2311	635.7078	31.5021
KPI12	7.5817	-0.0001	0.9063	0.0000	-0.0031	1307.8147	1319.1196	1.8499
KPI13	3935142.5931	-198.9862	0.7821	0.0276	-0.2965	111.2290	110.0573	8990.8659
KPI14	19541.0385	-1.0015	0.1281	0.5252	0.3669	40.0027	38.8311	7.2524
KPI15	-242415790.0729	12597.7796	0.4686	0.4065	-0.1870	85.7359	83.0318	142969.4002

Table 15: Linear regression results of performance by date of observation before commencing the project by KPI

What's immediately apparent by a quick examination is that the p-values are notably high across nearly all observations, with none of them falling below the commonly accepted threshold of 5%. This indicates a failure to reject the null hypothesis, suggesting no significant correlation as the coefficient of the variable *date_obs*. However, the most shocking observation lies in the values of R-squared (R2) and adjusted R-squared (R2 Adjust).

As depicted in the summarizing table below, which presents the means and standard deviations of the indicated statistics derived from the linear regressions:

After	p-value		R2		R2 Adjust	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
No	0.5974	0.2161	0.0923	0.1378	-0.0201	0.1311
Yes	0.1523	0.2829	0.2054	0.1728	0.1538	0.1509

Table 16: Mean, std. deviation for p-value, R^2, and R^2 adjusted of linear regressions of performance by date of observations before and after commencing the projects

The most relevant finding is the negative mean of the adjusted R^2, implying that using the mean value for predicted values would yield more accuracy than employing the actual linear prediction. This phenomenon could potentially be attributed to notably negative values observed for KPI13 and KPI15, which drag down the overall mean of adjusted R^2.

Generally speaking, the results align with initial expectations, indicating no discernible correlation between data and performance, which confirms that data might be randomly distributed around the mean.

Upon closer examination of the preceding table, we might ask whether the shift in the mean of the p-value, coupled with a relatively stable standard deviation and a doubled R^2, truly means an improved correlation between dates and performance, which might potentially suggest that consultancy efforts have yielded the desired effect.

Nonetheless, after closer examination and recalling earlier considerations regarding target benchmarks relative to the baselines, we can proceed to evaluate the results of regressions involving only post-project commencement observations. The following paragraphs outlines these results.

KPI id	Target Higher	Intercept	Coeff. Date Observation	p-value	R2	R2 Adjust	AIC	BIC	RMSE
KPI1	NO	22.9452	-0.0009	0.7750	0.0056	-0.0607	30.9808	33.4805	0.5045
KPI2	NO	746.8741	-0.0378	0.0003	0.2540	0.2371	213.4557	218.9416	2.3072
KPI3	YES	28.4920	-0.0014	0.2082	0.0963	0.0398	-22.2046	-19.5335	0.1105
KPI4	YES	8.5887	-0.0004	0.0011	0.1282	0.1170	-291.9592	-284.8131	0.0376
KPI5	NO	10.8528	-0.0005	0.0584	0.1662	0.1245	-56.9187	-53.6456	0.0579
KPI6	NO	5.0498	-0.0003	0.3853	0.0377	-0.0104	-55.2875	-52.0144	0.0601
KPI7	NO	222.5748	-0.0111	0.0011	0.4109	0.3815	47.9465	51.2196	0.6278
KPI8	NO	145.9505	-0.0072	0.0019	0.3824	0.3515	31.5134	34.7865	0.4321
KPI9	YES	1006.6775	-0.0504	1.56E-05	0.1527	0.1452	731.0106	739.2454	5.6592
KPI11	NO	8163.8522	-0.4104	3.59E-07	0.2053	0.1983	1172.1572	1180.3920	38.5251
KPI12	NO	466.8778	-0.0237	0.2564	0.0356	0.0088	164.7530	169.6657	1.9541
KPI13	NO	2814836.4785	-141.9372	0.1598	0.3309	0.1971	151.5775	151.4152	7938.3094
KPI14	NO	258.1286	-0.0131	0.1874	0.2992	0.1590	22.5185	22.3562	0.7874
KPI15	NO	236477202.9733	-11951.8098	0.0972	0.3702	0.2653	243.4304	243.6687	673647.9175

Table 17: Linear regression results of performance by date of observation after commencing the project by KPI

An initial observation reveals that all coefficients of the variable *date_obs* are negative, which immediately casts doubt on our model's ability to clarify diverse performance trends, in particular to identify the distinction between those variables that should trend upwards and those who should record a decrement. Furthermore, while there is a general decrease in p-values, some instances still exhibit significantly high values, particularly KPI1, KPI3, KPI6, and KPI11. Moreover, certain KPIs, distinctly KPI4 and KPI9, display concerning trends.

Despite expectations of positive coefficients for KPI4 and KPI9, our linear model yields negative coefficients, with related p-values indicating a strong correlation. But without control variables for comparison, it is difficult to elucidate the accuracy of this correlation.

In addition, it's crucial to note that although KPI4 displays a negative coefficient, this is very close to zero, meanwhile KPI9 exhibits a coefficient of only -5%, implying a performance decrease of 0.05 per additional day.

To gain deeper insights, let's explore their graphical representations:

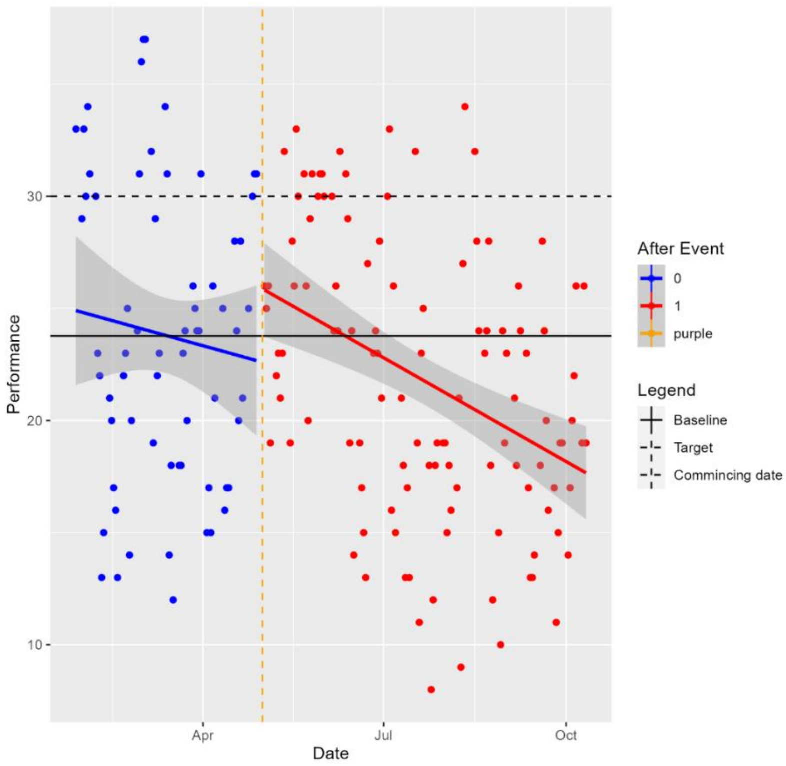

Figure 19: Scatter plot of performance by date of observation with linear tendency for before and after for KPI9

The recorded values reveal that while the *target*, indicated by the dashed upper line, is occasionally met, or exceeded, more often than not, it is not met. Instead, performance appears to be trending lower, indicating potential ineffectiveness of the consultancy efforts, and or the potential influence of unknown external factors.

Similarly, upon examining KPI4, which is here below, we note a negative performance trend. Nevertheless, without historical data, we cannot compare trend shifts, which force us to interpret data with an incomplete set of necessary information.

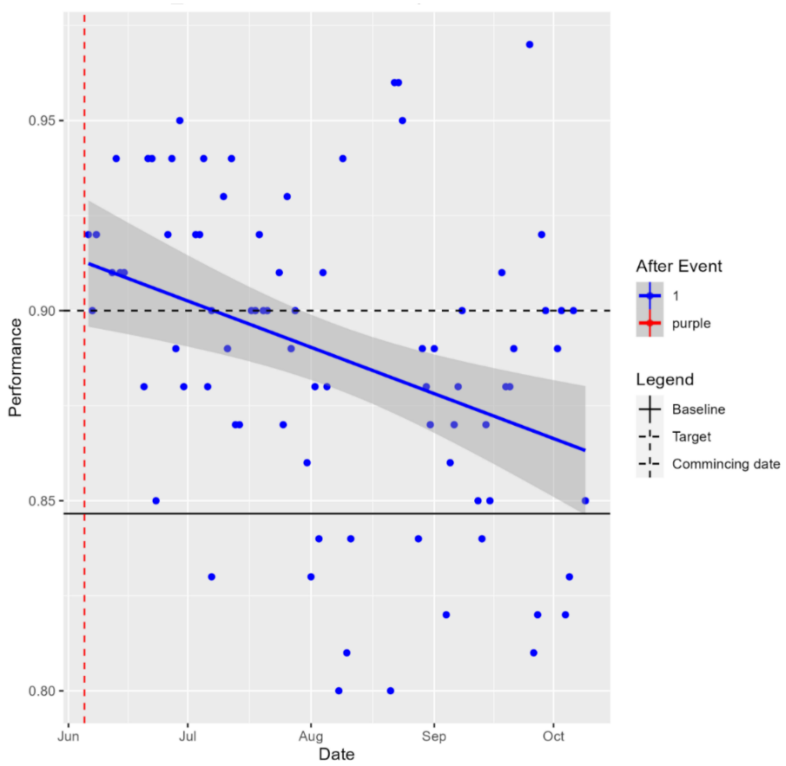

Figure 20: Scatter plot of performance by date of observation with linear tendency for after for KPI4

Some other notable performance can be observed if we take in consideration the following table, which tries to assess the value of the change between the coefficients and p-values, before and after the beginning of the project:

KPI id	Target Higher	Diff. Coeff.	Diff. Coeff. %	Diff. P-value
KPI1	NO	-0.0001	14%	-0.0699
KPI2	NO	-0.0441	-706%	-0.4103
KPI3	YES	-0.0015	-1356%	-0.6450
KPI4	YES	-0.0004	/	0.0011
KPI5	NO	-0.0007	-419%	-0.6502
KPI6	NO	0.0005	-67%	0.2723
KPI7	NO	-0.0136	-551%	-0.4116
KPI8	NO	-0.0059	444%	-0.8137
KPI9	YES	-0.0259	106%	-0.4427
KPI11	NO	-0.4335	-1884%	-0.8801
KPI12	NO	-0.0236	28269%	-0.6498
KPI13	NO	57.0490	-29%	-0.6222
KPI14	NO	0.9885	-99%	0.0593
KPI15	NO	-24549.5894	-195%	-0.3714

Table 18:Difference between coefficients (absolute and relative) and p-value before and after the commencing of the project by KPI

After careful consideration of the preceding table, it becomes evident that KPI2 and KPI11 exhibited the most significant and concordance change in their coefficient, in relative terms, while also improving their p-value. However, a crucial distinction must be made.

KPI2 appears to adhere to a linear prediction with some variability, as indicated by the grey area representing the 95% confidence interval of all values. On the other hand, KPI11 follows a nonlinear trajectory, which should be the subject of further analysis in subsequent sections.

Here they follows:

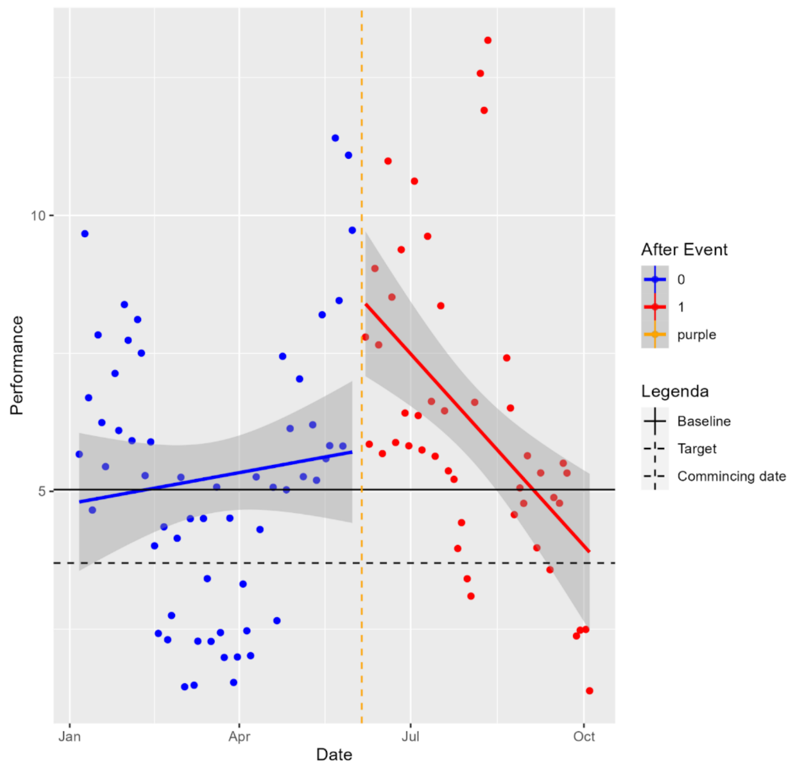

Figure 21: Scatter plot of performance by date of observation with linear tendency for before and after for KPI2

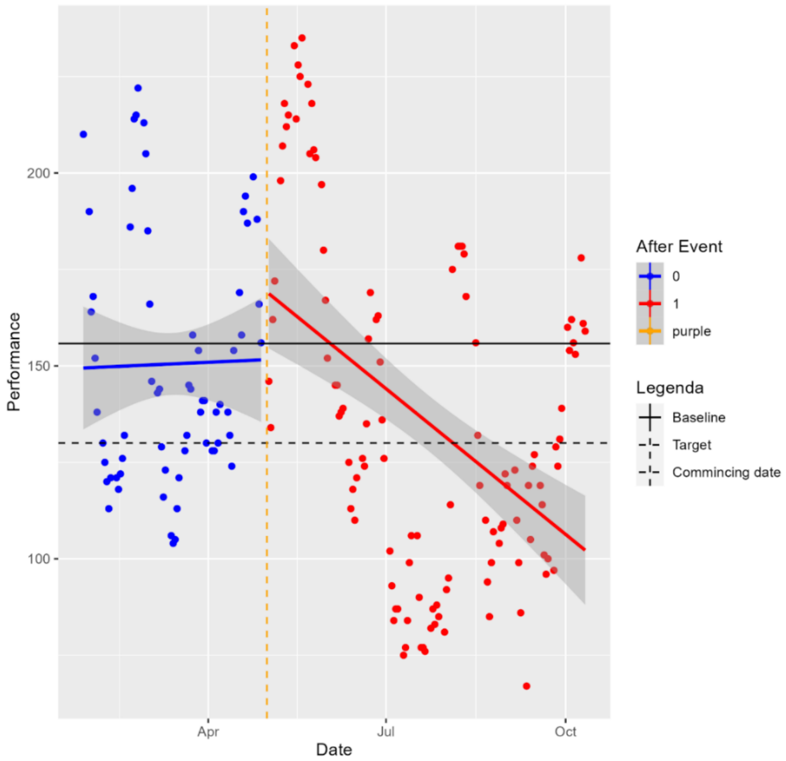

Figure 22: Scatter plot of performance by date of observation with linear tendency for before and after for KPI11

What a manager can glean from the preceding analysis is that a tangible impact has indeed been documented since projects' inception. Consequently, the activities undertaken, if not influenced by external factors, can be attributed to fischer Consulting's methodologies and interventions. However, numerous specific conditions necessitate further investigation. This is why, while it is undoubtedly a positive step towards the desired outcome, the concerns of a consulting manager should regard the solidity of the data and the variability inherent in the analysis of individual KPIs. Nevertheless, the uncovered information offers valuable insights into areas specific to the project and clients that require attention.

Furthermore, some of the observed trends are cause for pride as they demonstrate the additional value that fischer Consulting has contributed to the company's performances since its involvement.

4.4.2. Linear trends by area

The next step in this analysis involves assessing the impact on each KPI area involved in the study. Continuing with a similar approach to before, we initiated by conducting regressions for each distinct area, nevertheless some change in the variables involved was made. Given the unique nature of each KPI, it is impractical to regress performances of the same area against each other or against those of other areas, this is due to difference in units of measures and absolute values of these observations. Instead, what proves beneficial is evaluating these diverse measures using the unified metric, which is the target achievement, or TA.

An additional challenge was faced due to the diverse project start dates, which needed to be addressed, because since projects began at different times, using the date of observation as a regressor could potentially lead to different expected trends across various KPIs for the same date. However, by categorizing the date of observations as 'before' or 'after' project commencement we can completely avoid this complexity. So it was decided to opt to utilize *diff_since_beg*, which represents the difference in days between the recording date and the project's actual start date. Nonetheless, it is important to underline that none prohibit us to use date since the target achievement should be stationary before the commencing of the project, but this would give different weights to the date rather than the distance from the beginning of the project.

Consequently, following a similar methodology as before, the data were subsequently aggregated by category, and the results of the regressions are presented in the tables below. The first table summarizes the findings for the period before the projects' initiation. Here are the results:

KPI Area	Intercept	Coeff. Days since Beginning	p-value	R2	R2 Adjust	AIC	BIC	RMSE
QUALITY	-10,2021	0,0005	0,8449	0,0020	-0,0479	40,2069	43,4800	0,5265
EFFICIENCY	-1,3990	0,0001	0,8694	0,0001	-0,0019	1864,4321	1877,135	1,4964
PRODUCTIVITY	54,9458	-0,0028	0,4683	0,0047	-0,0042	450,3699	458,5520	1,7286

Table 19: Linear regression results of target achievement by difference from beginning of the project before commencing the project

By glancing at the table, what stands out are the p-values, which are consistently above the significance level, and this implies a failure to reject the null hypothesis. Furthermore, we can notice that the coefficient of the difference in days since the beginning is often near zero, as is the R^2 value. Moreover, the adjusted R^2 is negative, once again this suggests that using the mean would yield better predictions. This is in line with the expectations that we should not observe any particular trend to be persisting when evaluating data before an intervention by the consulting firm takes place, but we should look at values which vary randomly.

A crucial consideration that must be addressed concerns the structure of the TA. More specifically, it's tied to the fact that in the case of an undefined, by default, baseline, the mean value before the beginning of the project was chosen to be it. Therefore, when regressing the TA before and after the intervention, it's pivotal to remember that this is biased to an unknown extent, because those same data which were used to compute the *baseline* are then influenced by it when computing the TA. Consequently, we can expect all the TA observations recorded before to tend to gravitate towards the mean, which results in a flat trend line. While this was a necessary condition to assess whether there was a significant impact or not, it was underpinning in comparing the categories across multiple KPIs, given their different units of measurement.

Turning to the table summarizing the results for observations after the intervention by fischer Consulting Italia, several noteworthy observations emerge:

KPI Area	Intercept	Coeff. Days since Beginning	p-value	R2	R2 Adjust	AIC	BIC	RMSE
QUALITY	0,9885	-0,0060	0,0048	0,0805	0,0708	226,3180	234,0421	0,7533
EFFICIENCY	-0,4988	0,0144	3,29E-12	0,1567	0,1537	1075,599	1086,577	1,5597
PRODUCTIVITY	0,4369	-0,0073	2,28E-05	0,1109	0,1051	429,0501	438,1804	0,9472

Table 20: Linear regression results of target achievement by difference from beginning of the project after commencing the project

First and foremost, a significant shift in the p-values for each area has occurred, accompanied by a significant variation in the intercept. Particularly, the significance values of the p statistics are considerably low, indicating statistical significance, with the efficiency and productivity category below the significance level. Additionally, there has been a notable change in both the R^2 value and its adjusted counterpart, which are now both positive and hovering around 10%. One particularly notable aspect is the value of the coefficients, which are practically zero for quality and productivity, with the latter even exhibiting negative values, suggesting that the impact of the consultancy has not yielded the desired effects.

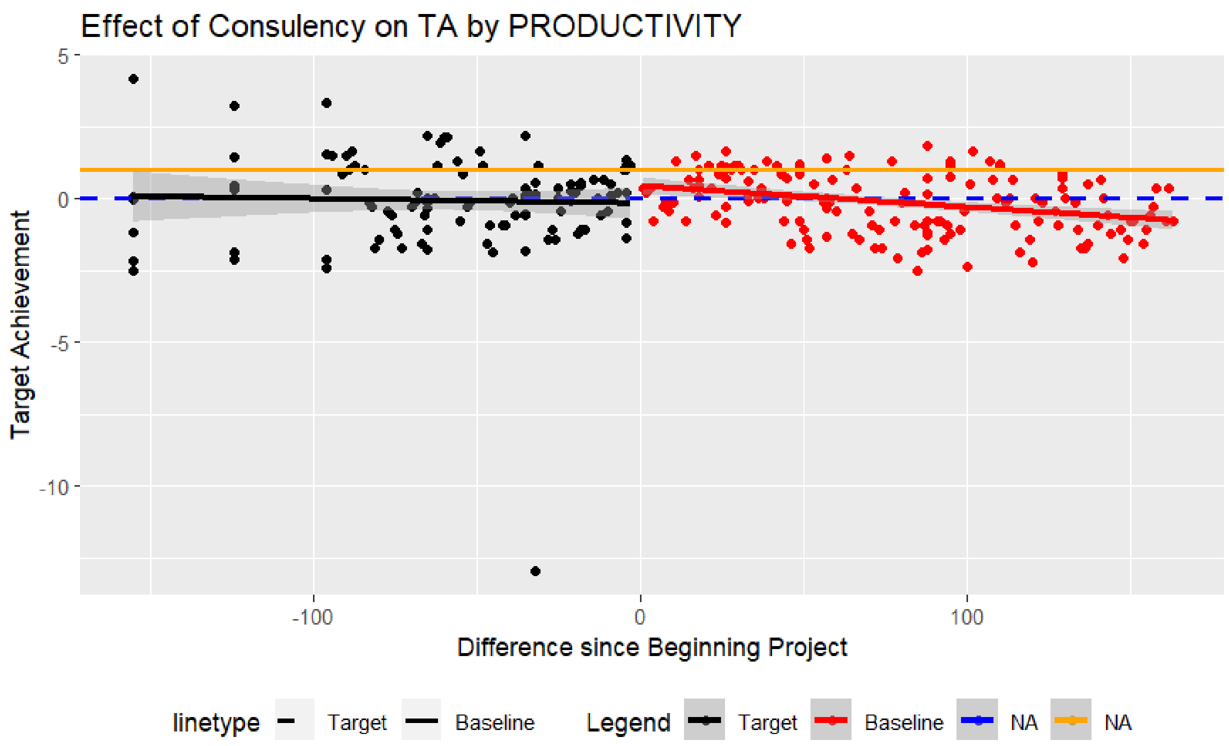

Figure 23: Scatter plot of target achievement by difference from beginning of the project with linear tendency for before and after for productivity

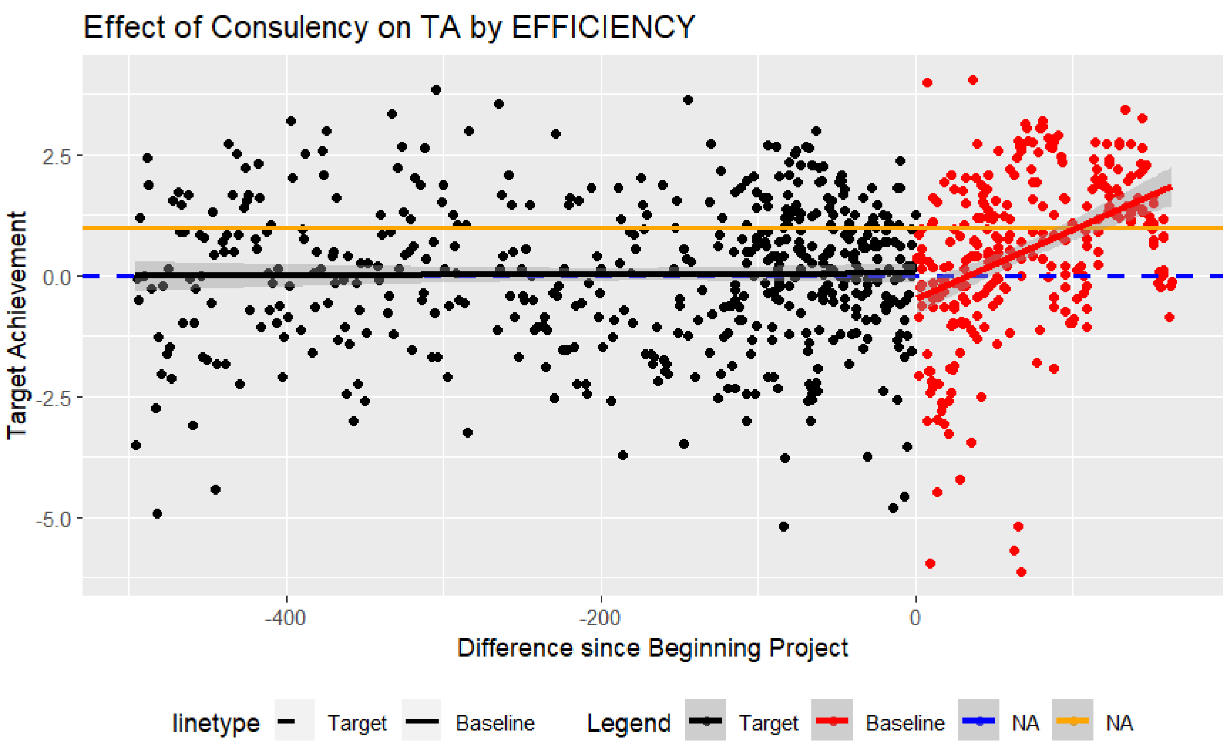

Figure 24: Scatter plot of target achievement by difference from beginning of the project with linear tendency for before and after for efficiency

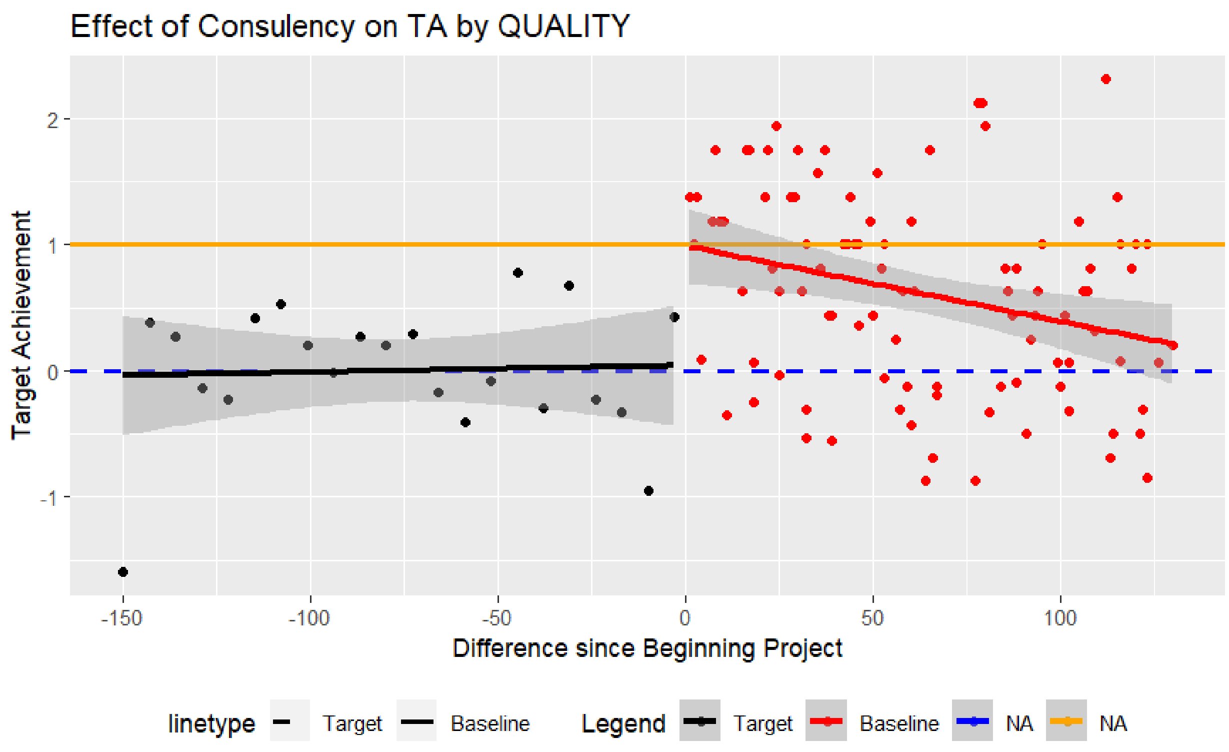

Figure 25: Scatter plot of target achievement by difference from beginning of the project with linear tendency for before and after for quality

After having examined the graphical representations of the regressions, we can easily spot that the confidence interval is wider for quality compared to the other areas. Unfortunately, both quality and productivity appear to exhibit downward slopes, with the former showing a steeper decline than all three areas. However, the most striking result is related to efficiency which demonstrates a clear and defined change in the tendency following the commencement of the project, regardless of its significant variability.

Once again, these information prove vital for a consulting firm manager when targeting specific projects, since they furnish essential insights into the team's skillset and their efficiency across distinct areas of expertise. Armed with this knowledge, the manager can make informed decisions, steering away from projects ill-suited to the team's strengths. Moreover, it opens the way for enhancing the team's existing skills or introducing experts with proven track records within the team. Such experts can share valuable insights, guiding the team toward achieving desired goals effectively and efficiently.

4.4.3. Linear tendency by KPI, a more comprehensive analysis

In this third subsection of the linear trend analysis, we delve into a more intricate model of linear regression. This model not only includes the date of observation but also considers the time period and the interaction between them.

The structure of this regression is designed to verify whether there are statistically significant relationships, particularly examining the impact of the variable *after* and its interaction with the date of observation, regarding, obviously, the prediction of the recorded performance.

The results of these regressions, conducted at the individual KPI level, are summarized in attachment number four, located at the end of this document due to spatial constraints. By contrast, the subsequent are summarizing results:

	P-Value			R2	R2 Adjust
	Date of Obs.	After	Interaction		
Mean	0.6369	0.3333	0.3340	0.2694	0.2034
Std. Deviation	0.3027	0.3298	0.3304	0.2161	0.1768

Table 21: Mean and std. deviation of regression results of performance by date of observation, after and their interaction by KPI id

Upon examination, the overall assessment of our regression is for the majority negative. Even incorporating the control variable *after* and its interaction with the date of observation fails to reject the hypothesis of a coefficient of zero for all three variables, at a general level. However, a deeper analysis is needed, as suggested by the considerable variability in p-values across variables and KPIs.

A closer look reveals a frequent pattern: most p-values indicate statistically insignificant coefficients, making inference impossible. Furthermore, KPI4, with values recorded only after project commencement, lacks coefficients for the *after* variable or its interaction, hence the value *NA*, which is the abbreviation for not available.

If we turn our attention to KPIs that demonstrate statistically significant values post-analysis, we identify KPIs 2, 7, 11, and 14.

It's crucial to elucidate a distinction: apart from KPI14, all others exhibit insignificant p-values for the date of observation, including numbers 2, 7, and 11, so we must keep in mind this when evaluating the individual statistics. Now, let's see the graphical representation for KPI14, elaborated in subsection 4.4.1. and 4.4.2.

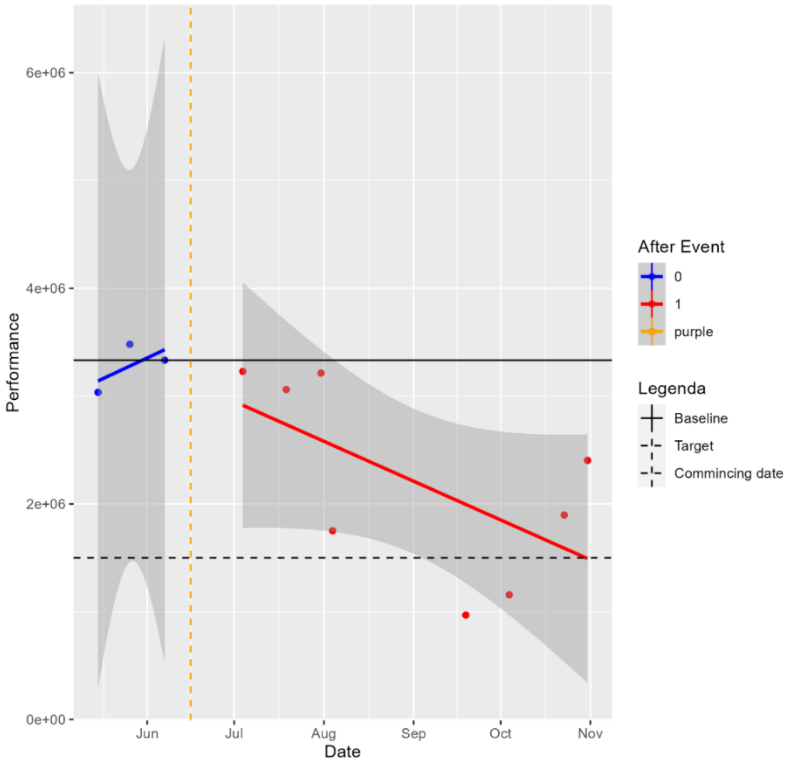

Figure 26: Scatter plot of performance by date of observation with linear tendency for before and after for KPI14

The first and most evident detail is the limited number of data points. We have a total of only eleven observations, with three occurring before and eight after. Despite the clear negative trend, what stands out are the 95% confidence intervals, in gray. Notably, these intervals are so extended that the lower limit of values before could reach the actual target level merely due to chances. To understand whether this variability persists in the new regression, let's examine the standard errors and other metrics. Below are the results:

Name	Value
(Intercept)	−242415790.073
	-870613475.2
date_obs	12597.78
	-44639.209
after	478892993
	-878366963.1
date_obs × after	−24549.589
	-45032.819
Num.Obs.	11
R2	0.559
R2 Adj.	0.371
AIC	333.1
BIC	335.1
Log.Lik.	−161.574
F	2.964
RMSE	579320.36

Table 22: Table results for the regression of performance against date_obs, after and their interaction, for KPI14

Our concerns are unfortunately validated, as indicated by the standard errors, which are values beneath the coefficient values of variables and their interactions. In particular, these values are significantly high in all three cases, exceeding even their own coefficients. This discrepancy underscores that the variability in performance values could be, possibly, twice as much as what the model predicts.

Moving now to other KPIs, let's consider KPI2, where the results for the p-value align closely with the standard error predicted by the model. Here the specific numbers:

Name	Value
(Intercept)	7.582
	-13.736
date_obs	0
	-0.001
after	459.296
	-373.856
date_obs × after	−0.024
	-0.019
Num.Obs.	358
R2	0.021
R2 Adj.	0.013
AIC	1470.8
BIC	1490.2
Log.Lik.	−730.389
F	2.584
RMSE	1.86

Table 23: Table results for the regression of performance against date_obs, after and their interaction, for KPI2

Specifically, the R^2 value is notably low compared to other non-significant KPIs, which raises some concerns. However, the regression appears consistent, as evidenced by attachment four, where the coefficient of the date_obs is insignificant, with a value of 0.006 and a standard error of 0.007, exceeding 100%.

KPI11 and KPI7 exhibit similar patterns to KPI2, as evidenced by their p-values. Particularly, KPI7 is still affected, much like KPI14, by a limited dataset, potentially impairing the precision of the overall model. Here's a graphical representation for clarity:

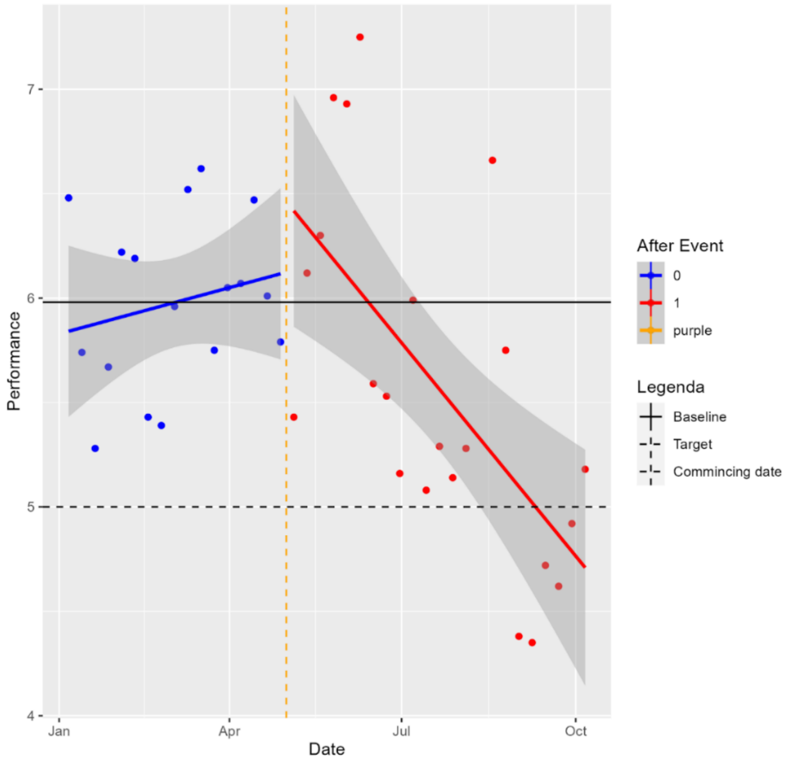

Figure 27: Scatter plot of performance by date of observation with linear tendency for before and after for KPI7

As a manager, these results are indeed concerning, as they suggest a weak correlation between actual performances and the variable data thereafter. However, it's important to note that while this analysis has indicated non-significance, it doesn't definitively rule out a causal relationship. Rather, it suggests that at the commonly accepted 5% alpha level, there's insufficient evidence to reject our null hypothesis, implying that observed changes may not be attributed to the variables under consideration, but might be simply due to chances.

4.4.4. Residuals Analysis

In precedent subsection we have delved into two crucial concepts: heteroskedasticity and homoscedasticity. These features provide valuable insights into the distribution of errors, illustrating whether they adhere to a normal distribution around a mean, are uniformly distributed, or follow any specific pattern.

The primary focus of our residuals analysis relies on the latest regression, which has considered the variables date_obs and after, along with their interaction. The resulting residuals are outlined in the table below:

KPI id	Min	1st_Quartile	Median	Mean	3rd_Quartile	Max
KPI1	-1,2175	-0,4740	0,0085	-1,55E-16	0,3596	2,4710
KPI2	-3,7844	-1,6288	-0,3178	1,05E-16	1,1036	7,2359
KPI3	-0,2670	-0,1175	-0,0024	-4,51E-18	0,0926	0,4655
KPI4	-0,0876	-0,0233	0,0024	7,16E-19	0,0245	0,1013
KPI5	-0,1318	-0,0505	0,0241	-4,44E-19	0,0522	0,0818
KPI6	-0,1064	-0,0491	-0,0038	1,66E-19	0,0420	0,1219
KPI7	-0,9878	-0,3450	-0,1184	-3,26E-17	0,3490	1,4072
KPI8	-1,0884	-0,3717	-0,0198	5,69E-18	0,2033	1,7265
KPI9	-13,5947	-4,4427	0,0457	5,25E-16	5,0653	13,2613
KPI11	-65,3981	-27,6613	-8,9131	1,72E-15	31,7959	74,9527
KPI12	-4,5720	-1,2307	-0,1340	-2,56E-17	1,2495	7,4672
KPI13	-14292,9200	-5533,1840	-1728,2590	1,29E-12	6003,2660	17294,4600
KPI14	-8,8457	-1,7147	-0,2300	-8,88E-16	0,9109	10,0250
KPI15	-1025907,0000	-382252,1000	202125,6000	1,85E-11	319360,7000	910874,5000

Table 24: Min, max, interquartiles, median, and mean of residuals of the regression of performance against date_obs, after and their interaction by KPI

As explained in previous chapters, the mean of the residuals should closely approximate zero, indicating that the minimization of the sum of squared residuals performed in a desirable way, even though the OLS disregard the sign of the residuals, due to its construction. However, it's imperative to pay close attention to the disparity between the median and mean recorded. Notably, three KPIs call for a careful scrutiny: KPI11, KPI13, and KPI15.

Taking a closer look at KPI13 and KPI15, which belong to the same client, we can visualize the distribution of residuals:

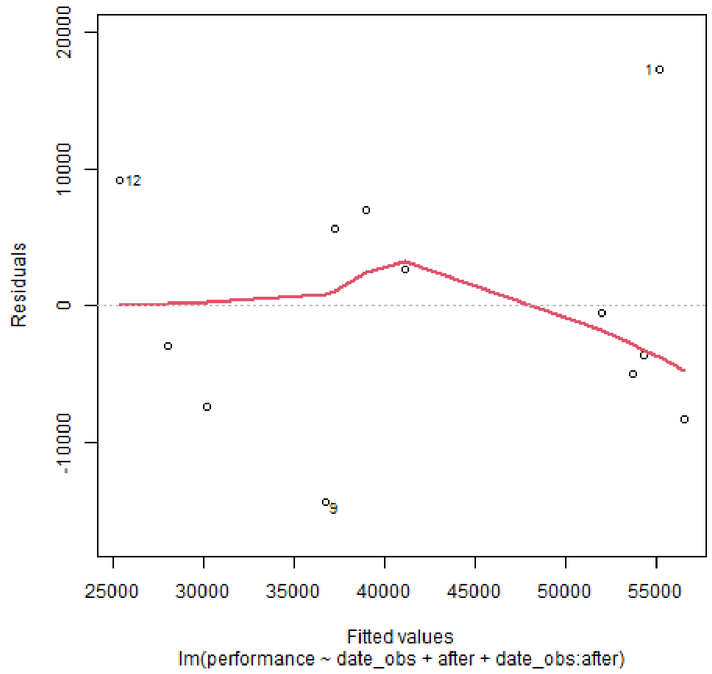

Figure 28: Residuals against fitted value of the regression of performance against date_obs, after and their interaction for KPI13

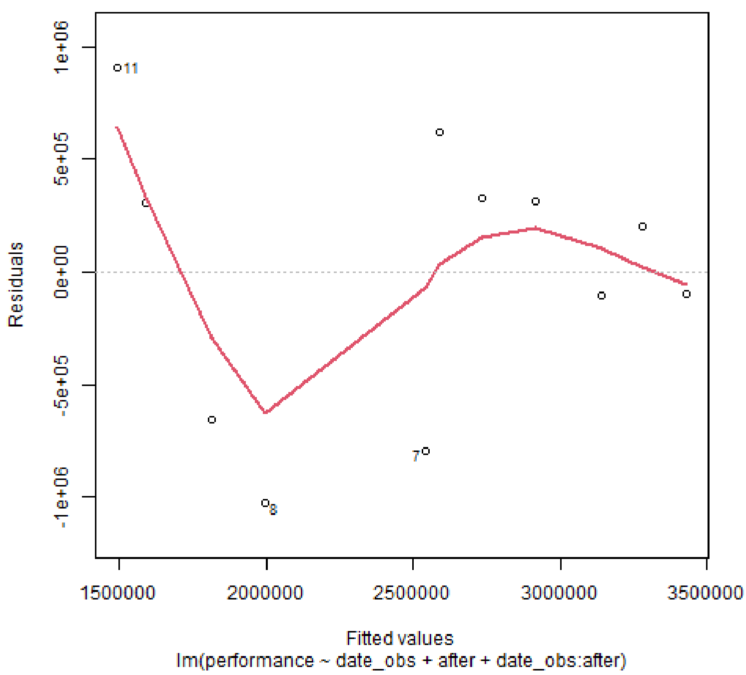

Figure 29: Residuals against fitted value of the regression of performance against date_obs, after and their interaction for KPI15

After an initial inspection, two distinct situations emerge: KPI13 doesn't raise significant concerns. Indeed, although its residuals exhibit considerable variability, as underlined by their distance from the horizontal zero, the trend line remains nearly flat, which is usually characteristic of homoscedasticity. On the other hand, KPI15 displays a widely fluctuating trend, largely influenced by a single factor affecting KPI13 as well: a limited dataset. It's clear that a handful of data points can influence disproportionately the direction of the trend line, particularly when there is significant variability.

We've previously examined KPI15 and are conscientious of its performance variability, it's important also to note that KPI13 also lacks fixed, or even stable, observations, as depicted in its trend line graph that follows:

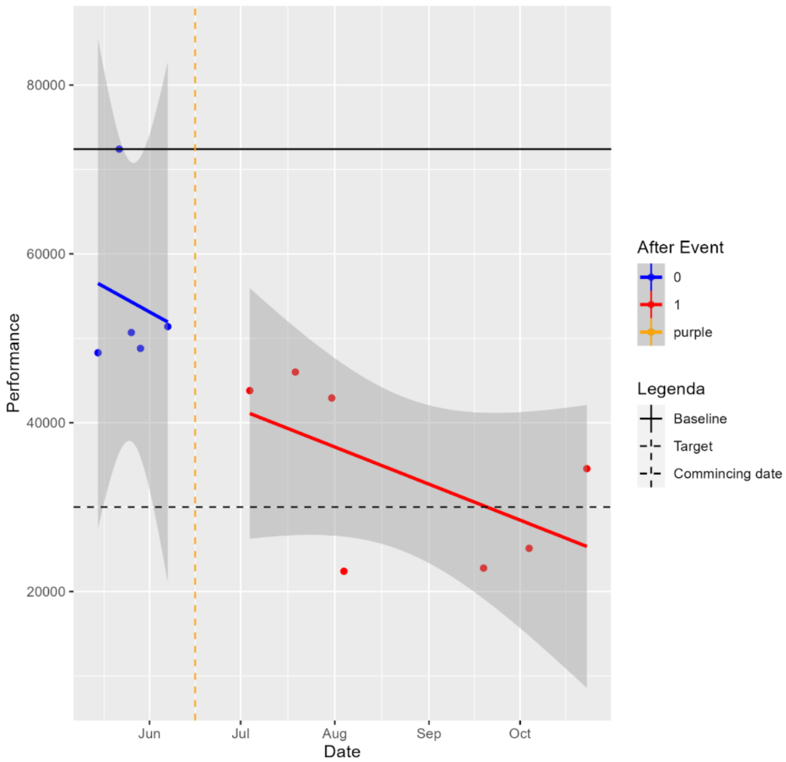

Figure 30: Scatter plot of performance by date of observation with linear tendency for before and after for KPI13

In light of this, it's prudent not to consider the disparity in residuals as a statistically significant factor for assuming either heteroskedasticity or homoscedasticity in the cases of KPI13 and KPI15.

However, KPI11 argues for a different story. Its residual graph exhibits a distinct peculiarity that demands closer attention:

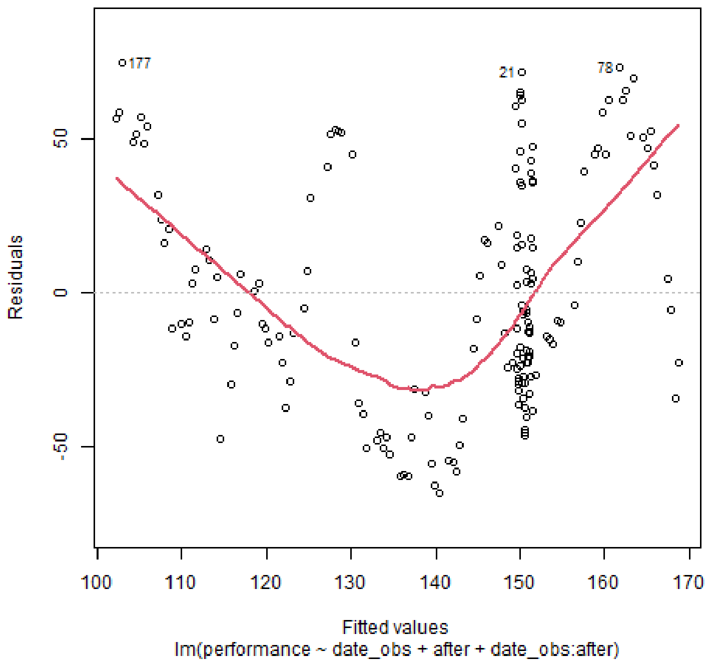

Figure 31: Residuals against fitted value of the regression of performance against date_obs, after and their interaction for KPI11

This KPI exhibits a significant level of discrepancy across various fitted values, specifically taking the form of a concave line with heterogeneous prediction errors. At both extremes and the center, it registers significantly higher errors compared to the two interquartiles of the fitted value range. While this is undoubtedly influenced by a substantial concentration of data around the fitted value of approximately 150, there's a relevant and consistent pattern of observing divergent errors. Even though this linear regression yielded statistically significant values for *after* and the interaction, it did not for *date_obs*. This raises doubts about whether this regressive model allows us to confirm or dismiss apparent heteroskedasticity. Nevertheless, it certainly provides reasons for caution concerning the potential presence of heteroskedasticity and requires further investigation into whether an alternative fitting line, different from a simple linear model, could better explain its performance movements.

Regarding the other KPIs, their residual distributions appear to follow a common pattern, with the trend line mostly laying around the zero or exhibiting irrelevant fluctuations away from this central value. The sole exception once more involves client 4 and is about KPI14, which is afflicted by a small number of data points. Below is its graph:

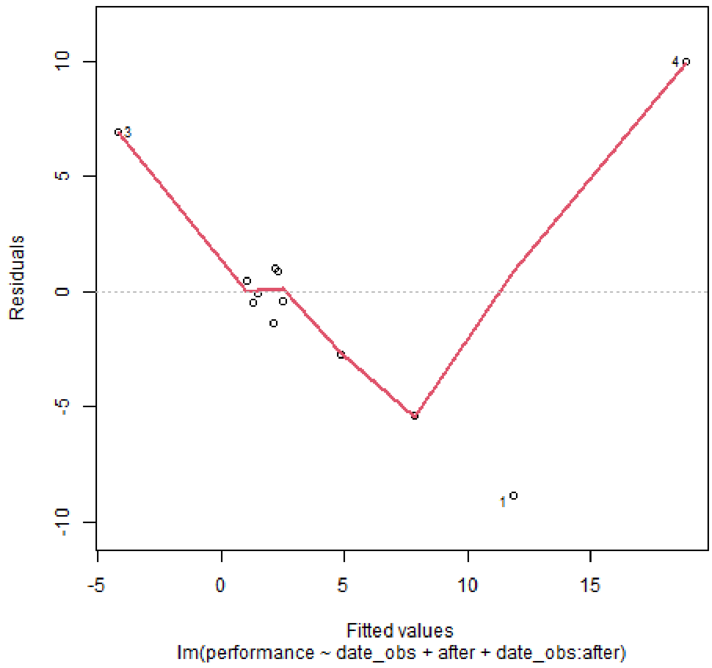

Figure 32: Residuals against fitted value of the regression of performance against date_obs, after and their interaction for KPI14

4.5. LOESS

An essential step in uncovering potential non-parametric hidden trends is to employ LOESS (Locally Weighted Scatterplot Smoothing) analysis. As discussed in Chapter Two, LOESS generates local smoothed lines, which are then aggregated to form the final trend curve. The results of such analysis are going to be discussed in the next pages.

4.5.1. Overview of results

During the analysis, two distinct LOESS regressions were conducted for each KPI. The former focused on data recorded prior to the beginning of the project, while the latter pertained to data recorded afterward. All the generated graphs contain a shaded gray area around the trend line, which represents the 95% confidence interval, like those observed in Section 4.4.

Due to the localized nature of LOESS regression, testing the significance level of coefficients, as done previously, was not feasible. This is because the span parameter, which determines smoothness and varies to fit current data, results in a different number of coefficients for each local area where the smoothness is executed. Consequently, the following table presents only the mean of the predicted values alongside the actual ones, along with their absolute and relative standard errors to the fitted values.

KPI id	After Yes=1 No=0	Mean Performance	Mean Fitted	Standard Error	% Std. Error
KPI1	0	4,0899	4,1036	0,8264	20,14%
KPI1	1	4,3662	4,3612	0,5962	13,67%
KPI2	0	5,2513	5,1897	1,5724	30,30%
KPI2	1	6,2692	6,2101	2,2706	36,56%
KPI3	0	0,5910	0,5901	0,1970	33,38%
KPI3	1	0,5989	0,5989	0,1203	20,08%
KPI4	1	0,8881	0,8875	0,0383	4,32%
KPI5	0	0,2435	0,2438	0,0637	26,15%
KPI5	1	0,1543	0,1538	0,0631	41,03%
KPI6	0	0,1954	0,1970	0,0638	32,39%
KPI6	1	0,1236	0,1250	0,0600	48,03%
KPI7	0	5,9788	5,9849	0,4276	7,14%
KPI7	1	5,5741	5,5505	0,6164	11,10%
KPI8	0	0,2435	0,2438	0,0637	26,15%
KPI8	1	0,1543	0,1538	0,0631	41,03%
KPI9	0	0,1954	0,1970	0,0638	32,39%
KPI9	1	0,1236	0,1250	0,0600	48,03%
KPI11	0	5,9788	5,9849	0,4276	7,14%
KPI11	1	135,3478	134,6796	29,2068	21,69%
KPI12	0	5,9774	5,9898	1,8427	30,76%
KPI12	1	5,1771	5,1592	2,0470	39,68%
KPI13	0	54314,6000	54314,6000	NaN	NaN
KPI13	1	33941,2857	33927,7735	9058,8271	26,70%
KPI14	0	7,8395	7,8395	NaN	NaN
KPI14	1	1,8579	1,8548	1,1330	61,09%
KPI15	0	3282904,6667	3282904,6667	NaN	NaN
KPI15	1	2209779,1250	2211205,2040	559075,9649	25,28%

Table 25: Mean performance and fitted values with standard errors of LOESS model, both relative and absolute, by KPI and after

The data in the table are categorized based on whether the data were recorded before or after the project commencement. Two notable observations emerge immediately. Firstly, there are some missing values, particularly evident in KPI13, KPI14, and KPI15. Secondly, there is considerable heterogeneity among the percentages, with some nearing or exceeding 50% while others remain notably low.

Given these considerations, it was proposed to divide the results into four arbitrary categories:

- Values with standard percentage errors below 30% during both periods.

- Values with standard percentage errors below 30% in only one period.
- Values with standard percentage errors above 30% during both periods.
- Instances where standard percentage errors are missing.

Before delving further, it's crucial to clarify an expectation: akin to linear regression, we anticipate a relatively flat trend for LOESS before the project's onset, with the 95% interval gradually narrowing as the project progresses and consultants implement their solutions.

Considering the nature of LOESS, we do not anticipate widely asymmetric distributions around the tendency line within the 95% interval. LOESS is non-parametric and structured to nonlinearly follow recorded data.

Beginning with condition (1), the following KPIs fall into this category: 1, 4, 7, 8, 9, and 11. These LOESS regressions typically exhibit most data points falling within or near the 95% intervals, resulting in low mean standard errors. This could be attributed to the close alignment of all performances. For instance, KPI1 exemplifies this trend:

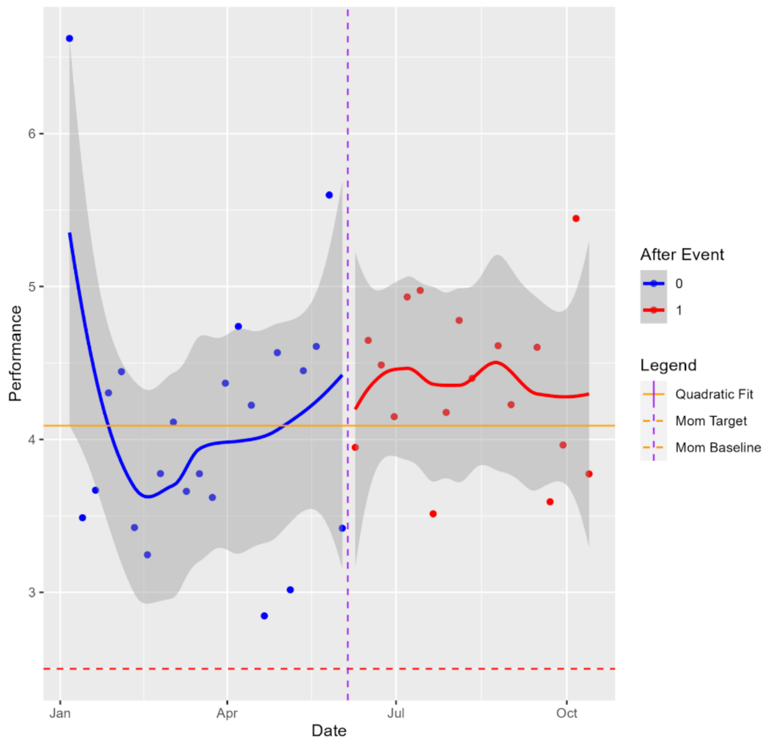

Figure 33: LOESS plot of performance by date of observation with non-parametric tendency for before and after for KPI1

The majority of data points fall within the interval, aligning closely with the underlying trend lines, with only the extremes showing greater divergence than the rest. This pattern is consistent across the other KPIs in this category, except for KPI11, which shows a notable amount of data lying outside this range, as evidenced by:

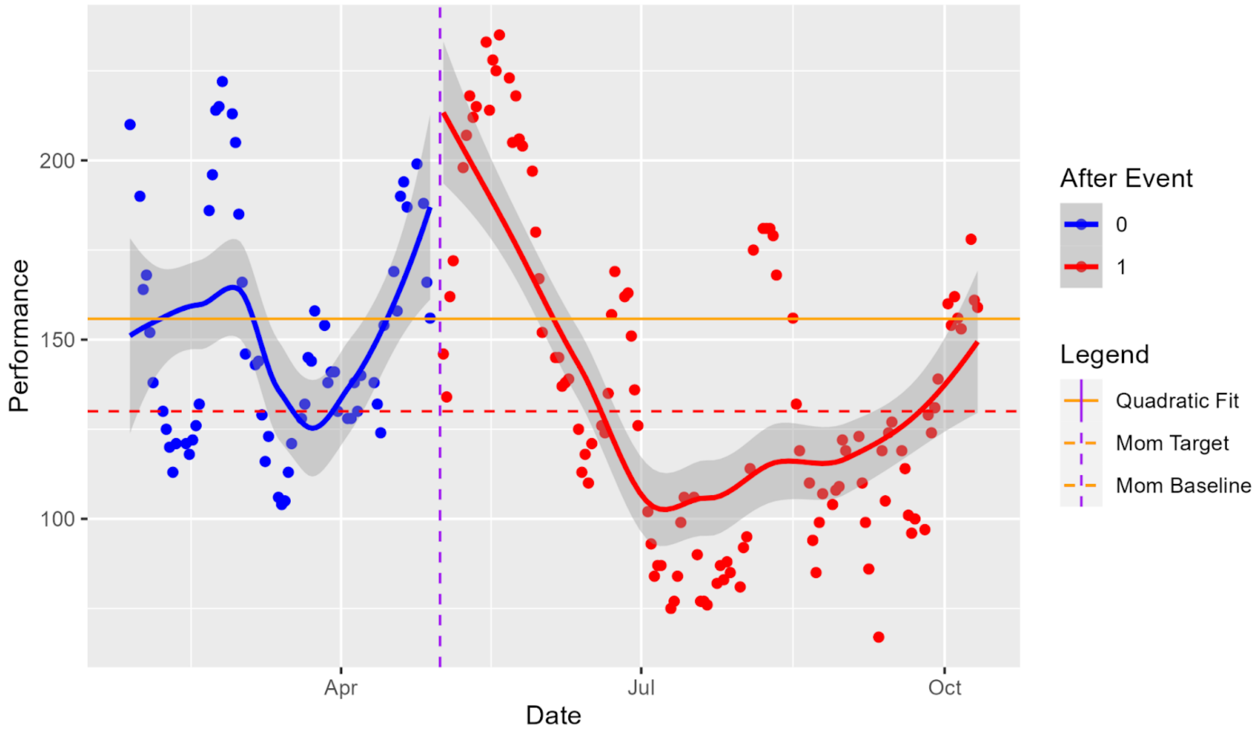

Figure 34: LOESS plot of performance by date of observation with non-parametric tendency for before and after for KPI11

It's worth noting that the errors as a percentage of the fitted value were 25% and 26% for the periods before and after, respectively. Additionally, the trend line does not adhere to a simple curve with occasional inversions of tendency; instead, it is characterized by a complex shape with multiple changes in direction. As discussed in the preceding chapters, this particular KPI demonstrates a unique trend, which is clearly nonlinear in nature. Nevertheless, despite the positive identification of the general trend by the loess analysis, it fails to fully explain a significant portion of the overall data due to the considerable variability within recorded performances.

Moving on, in condition (2), we observe KPI3 and KPI5. It follows the graphical representation of KPI3:

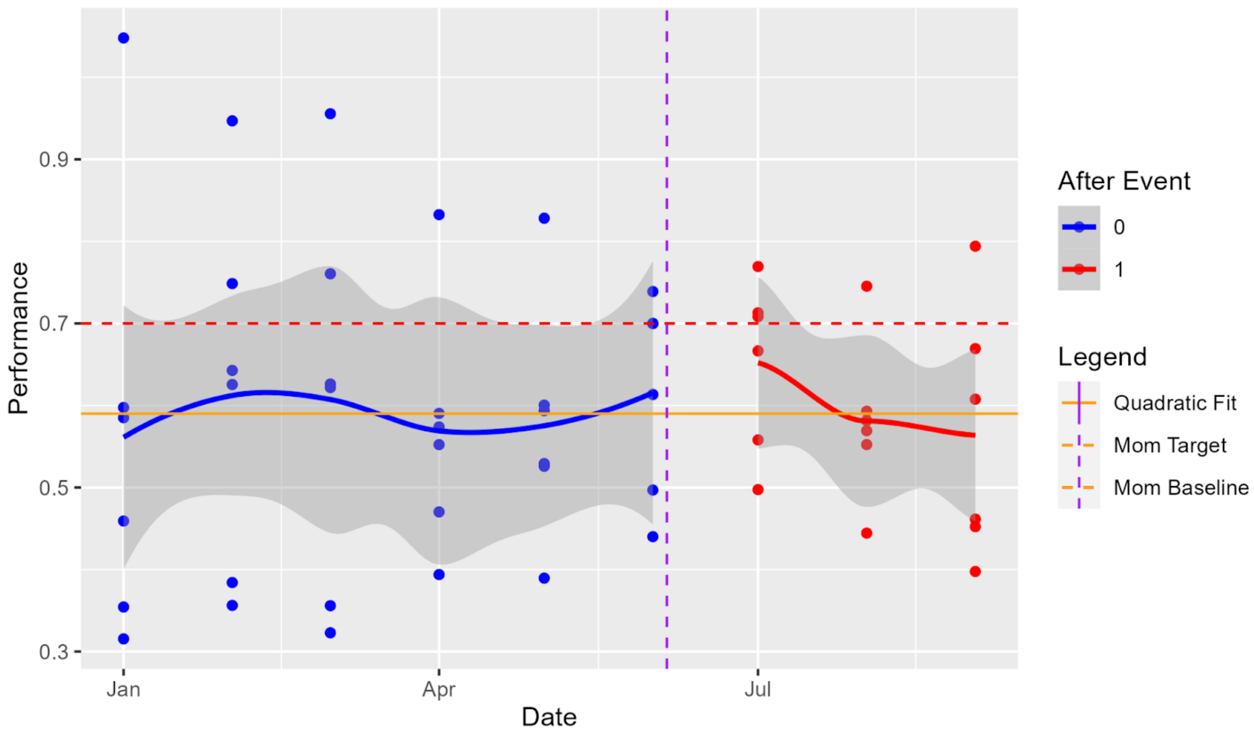

Figure 35: LOESS plot of performance by date of observation with non-parametric tendency for before and after for KPI3

As it is possible to see, before the project's inception, we observed a trend of sparser points which are drifting further away from the central line. This is evident also by looking at the narrowing of the gray area and the reduction in the range of performance, indicating a more pronounced similarity in performances. Something that must be remembered is that the data for KPI3 are aligned due to their weekly frequency of recording, as discussed in Chapter 3, despite being measured daily.

In conditions (3), KPIs 2, 6, and 12 exhibit notably high standard errors, suggesting that the LOESS model inadequately predicts and represents underlying data trends. For example, KPI12's LOESS model fails to capture a nonlinear trend.

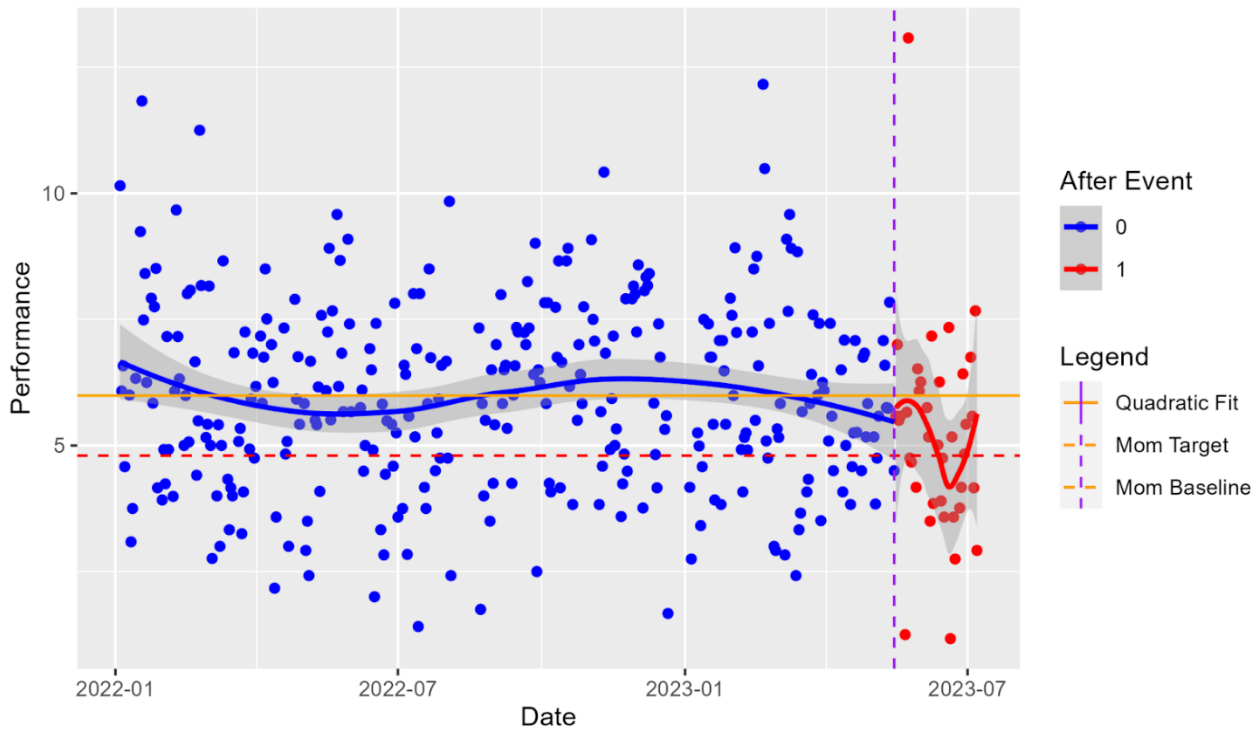

Figure 36: LOESS plot of performance by date of observation with non-parametric tendency for before and after for KPI12

The data are uniformly dispersed around the baseline, and the substantial number of observations highlights significant discrepancies between actual and predicted values. Conversely, a more pronounced curvature emerges in the post-period, indicative of a negative trend towards target achievement, as clarified in previous chapters.

Condition (4) regards exclusively KPIs 13, 14, and 15 of client 4. These are distinguished by their lack of a defined standard error. This absence is due to a zero span in the before period, indicating an exact fit to the data. This is related to the limited pre-project information available for these KPIs. Consequently, the local smoothness is confined to a single point, which necessitates by construction that the best-fitting curve will pass through that singular point, as is also possible to verify across all graphs. Consider the representation of KPI15:

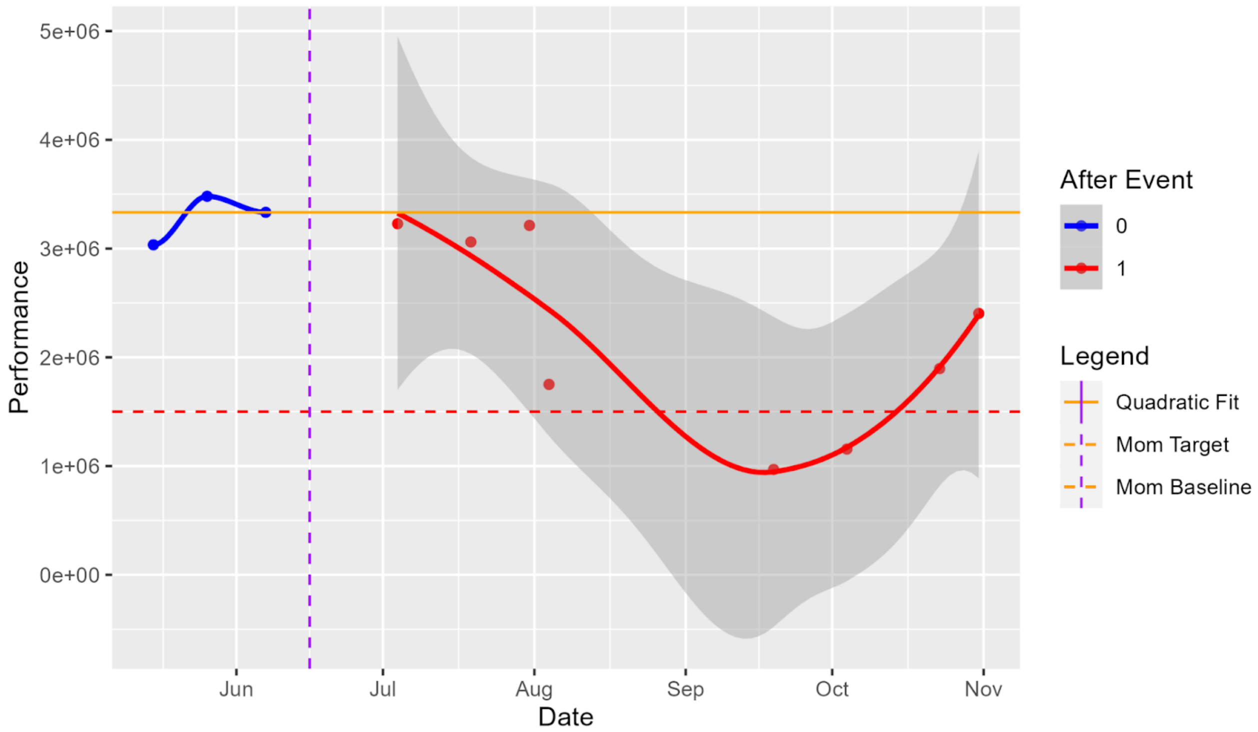

Figure 37: LOESS plot of performance by date of observation with non-parametric tendency for before and after for KPI15

Clearly, there is no confidence interval in the preceding period, as the local smoothness exclusively includes one point for each area for which the slope coefficient is computed. This pattern persists into the subsequent period, particularly the final months, where the red trend line precisely intersects the last four points. This consistency is also observable in the graphs for KPI13 and KPI14.

From a managerial standpoint, grasping the non-parametric performance of KPIs has significant importance for one primary reason: it enables managers to discern whether any improvement patterns, if present, adhere to linear and consistent projections, as might be suggested by linear regression results, or if they are better explained by a non-linear model. This understanding is crucial for identifying trends that may delay the attainment of desired targets.

For instance, conducting a LOESS analysis provides management with the opportunity to evaluate the timeframe within which the consulting team achieves optimal outcomes. If, for example, the consultants' intervention focuses in the *QUALITY* domain and consistently exhibit an initial downturn before rebounding to meet targets, then using a linear compensation model may overlook this pattern. By leveraging the modeled LOESS function, managers can optimize the compensation time frames to maximize both customer results and fischer Consulting's compensation. In this scenario, setting the checkpoint for variable compensation at the third month, when earnings could be potentially higher by the time we get to the fifth, would be inefficient.

Furthermore, having a track record to illustrate and reassure clients about potential initial performance reductions can prevent premature project shutdowns or sabotage due to temporary setbacks, because of the future expected benefits.

Moreover, this understanding allows both parties to better calibrate consultancy compensation, improving its alignment with performance and target achievements reached. It would be also possible to compute potential costs beforehand, such as increased defective products, and forecast future savings, which may result from improved quality methods.

4.5.2. LOESS trend by area

The next step involved evaluating the variation within different categories. Similar to the approach used in the linear analysis, a loess regression was conducted on target achievement both before and after the project commencement.

As previously, the difference since the project's beginning served as the regressor, enabling the minimization of the impact in differences across projects commencement date. The results are summarized in the following table, where "After" is represented as 0 for values recorded before the project and 1 for values recorded after its commencement.

Area	After Yes=1 No=0	Mean Actual	Mean Fitted	Standard Error	Std. Error %
QUALITY	0	0,0001	-0,0086	0,5197	-6056%
QUALITY	1	0,6109	0,5988	0,7681	128%
EFFICIENCY	0	0,0225	-0,0047	1,4847	-31394%
EFFICIENCY	1	0,5147	0,5122	1,5325	299%
PRODUCTIVITY	0	-0,0822	-0,0851	1,7373	-2042%
PRODUCTIVITY	1	-0,1351	-0,1623	0,9309	-574%

Table 26: Mean performance and fitted values with standard errors of LOESS model, both relative and absolute, by category and after

Upon initial inspection, the obvious aspect is the notable standard error in percentage proportions, which are significantly higher than those observed in the previous LOESS analysis. It's important to note that this is primarily due to the way target achievement (TA) is defined. Since the baseline, or the value at the project's beginning, should ideally be zero or close to it, the actual means for values recorded before the project are very close to this baseline. Consequently, since marginal differences exist between the actual and fitted means, the percentage of standard error gets inflated due to the very small denominator. Indeed, these large values are predominantly observed in observations recorded before the project. Conversely, examining lines where the *After* value equals one reveals that, while still relatively high, the actual percentage values have decreased significantly compared to before. For instance, the efficiency category has seen a reduction in percentage by a factor of 105, despite the actual change being only 32 times the original standard error.

Nonetheless, the percentage change in the "after" phase still indicates a significant standard error, which calls for caution when interpreting the overall results.

Beginning with QUALITY, let's delve into the evaluation of the current results:

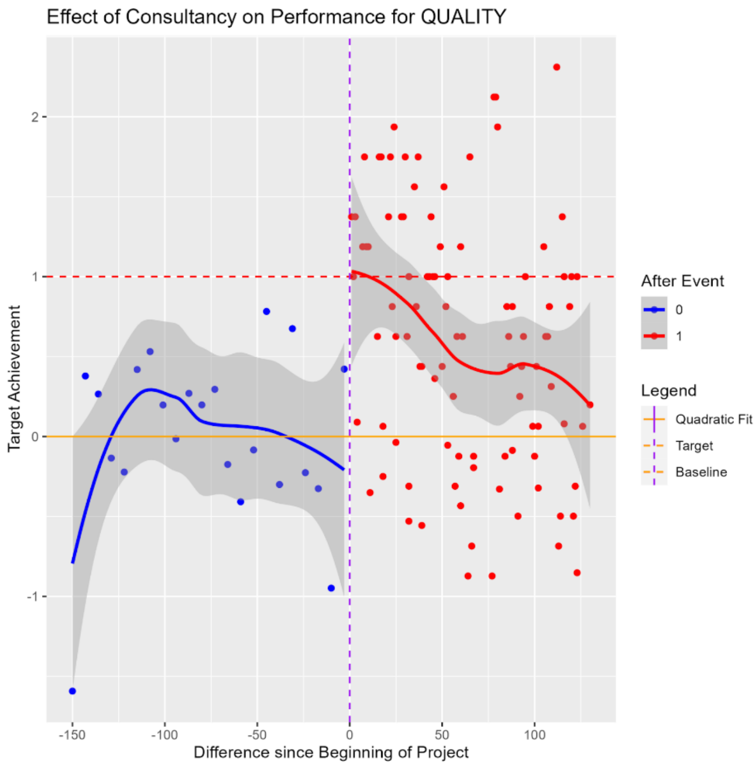

Figure 38: LOESS plot of target achievement by difference in days from beginning of the project with non-parametric tendency for before and after for quality category

The most noticeable observation is the concentration of data points predominantly on the right side of the zero difference in days, which outnumbered those with negative abscissa. Particularly intriguing is the width of the 95% confidence interval, since it is broader for observations before the project started, compared to those recorded afterward. What's particularly interesting is also the pronounced negative slope of the curve in the after period, indicating that while the consultancy initially achieved sudden and highly effective results, its effectiveness seemed to diminish over time.

A valid concern arises regarding the distribution of values across multiple KPIs, since more than one client KPIs' might be considered within a certain category. For instance, it is possible that the extreme concentration of observations on the far right may be driven by just one KPI, potentially skewing the overall trend of the line. In the case of the *QUALITY* category, only one KPI related to client one was observed, but this condition doesn't repeat for efficiency:

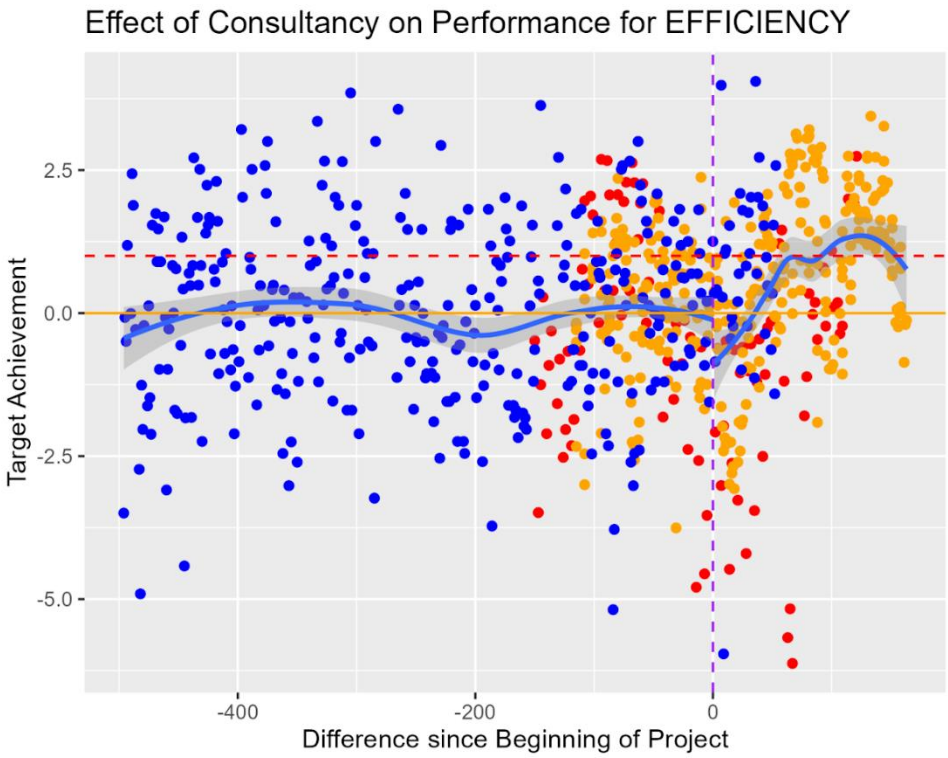

Figure 39: LOESS plot of target achievement difference in days from beginning of the project with non-parametric tendency for before and after for efficiency category

Subsequently to examination, it becomes evident that most values recorded 400 or more days before the project's commencement, highlighted in blue, are associated with a single client, namely client three. Conversely, observations after the project's commencement, highlighted in orange, are linked to client two, with the remaining red attributed to client one. While the graph still validates the overall trend observed for the category as a whole, it's important to acknowledge this disparity, as it addresses the possibility that data predictability may have been influenced by a single KPI exhibiting unpredictable or highly variable performance. While this scenario doesn't apply here, it's an important consideration to keep in mind. In the case of efficiency, a strong positive trend is evident, even surpassing the 100% threshold before slightly declining below it. Nonetheless, it serves as compelling evidence of fischer Consulting's effectiveness.

Moving on to the final category, productivity, we observe clients one, two, and three represented in blue, orange, and red, respectively. Once again, an asymmetry in the distribution of observations before and after the project's start is apparent. Interestingly, in this case, the LOESS trend remains relatively flat, even exhibiting a slight decrease in the second phase. Remarkably, a solitary red dot at the bottom center catches the eye, likely an outlier prompting questions about the accuracy of data entry.

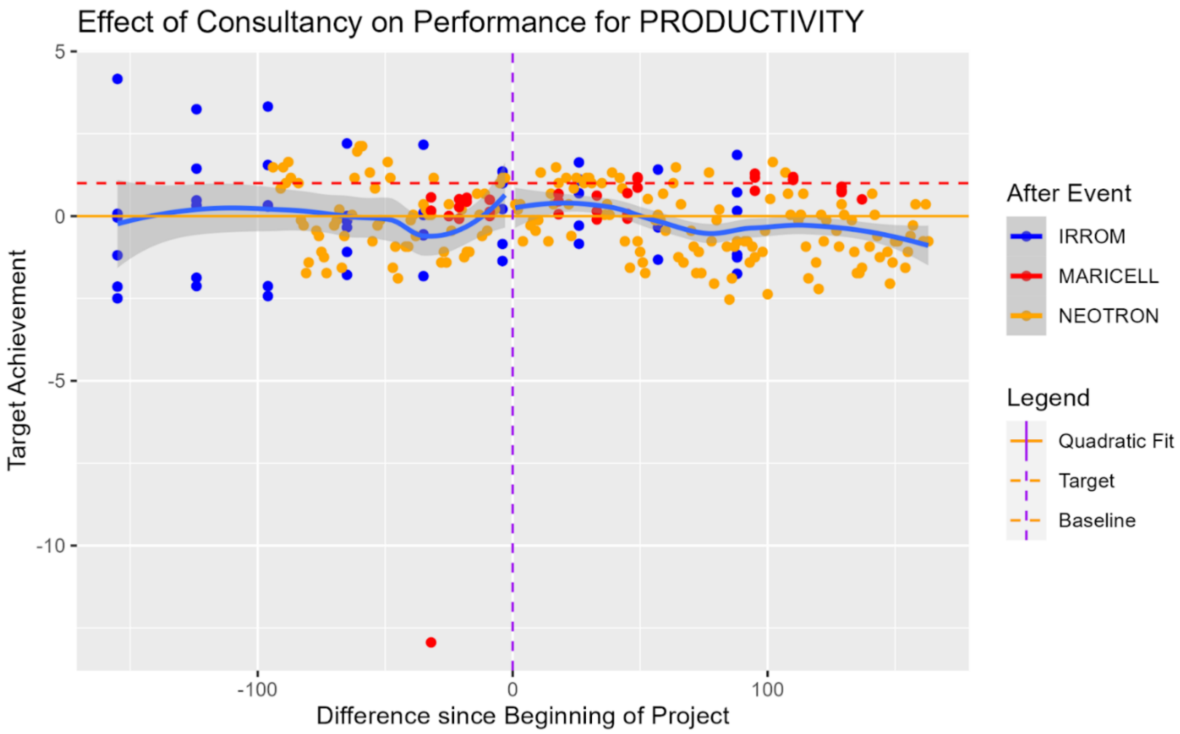

Figure 40: LOESS plot of target achievement difference in days from beginning of the project with non-parametric tendency for before and after for productivity category

As a manager, an understandable shadow of uncertainty is pending over the final conclusion about the interpretation of these results. While undoubtedly, there are discernible positive trends, as evidenced in the efficiency category, which clearly underline a strength of fischer Consulting; the remaining areas need some improvement, particularly in productivity and in the statistical significance of quality results. Addressing these specific issues is crucial for a comprehensive and effective evaluation of these two categories.

4.6. Time Series Analysis

In this subsection, we delve into the results of our time series analysis, conducted using pre-project data as training data for the time series, through the *ts( )* function. Later, we utilized the *auto.arima( )* function to forecast future movements and subsequently compared them with the actual data.

4.6.1. Overview of results

Our initial assessment focused on determining whether we had a sufficient number of observations to perform ARIMA analysis effectively. Regrettably, it was discovered that several observations fell short. Specifically, KPI numbers 4, 13, 14, 15, as well as those pertaining to CLIENT4 and one regarding client 1, lacked the minimum number of observations to support a viable ARIMA analysis. A minimum threshold of 10 observations was considered necessary in the preceding period for the computation of a time series model using ARIMA, essential for constructing predictive models based on this method.

The following table presents a summary of the ARIMA model results alongside their corresponding statistics:

KPI Id	AR & MA				Mean	Variance	Log. Likelihood	AIC	BIC
	AR1	AR2	MA1	MA2					
KPI1	NA	NA	NA	NA	4.0899	0.7356	-27.3271	58.6542	60.8363
KPI2	-0.6554	-0.7244	-0.2030	0.3757	NA	2.4444	-105.1483	220.2967	230.5120
KPI3	NA	NA	NA	NA	0.5910	0.0343	10.1553	-16.3106	-13.1436
KPI5	NA	NA	NA	NA	0.2435176	0.0039	23.6366	-43.2733	-41.6069
KPI6	NA	NA	NA	NA	0.1954412	0.0046	22.1145	-40.2291	-38.5626
KPI7	0.527985	NA	NA	NA	5.9914761	0.1420	-6.6294	19.2587	21.7584
KPI8	NA	NA	NA	NA	5.4505882	0.5877	-19.0890	42.1780	43.8445
KPI9	0.567881	NA	NA	NA	24.120581	30.8080	-199.6795	405.3591	411.8357
KPI11	1.845798	-0.9167	-0.7971	149.7174	NA	141.4464	-248.6623	507.3245	518.1189
KPI12	0.1296	0.110841	NA	NA	5.9809912	3.3407	-645.5636	1299.1273	1314.2005

Table 27: AR, MA, Mean, Variance, Log. Likelihood, AIC and BIC values for ARIMA model by KPI

As highlighted by the table, a significant number of *auto.arima( )* outputs returned *NA*, indicating their unavailability. This occurrence is inherent to the ARIMA methodology and suggests the absence of significant trends within the training data that could influence future trends. This could be attributed to several factors, one of which may be that the actual real trend is not afflicted by autoregressive and moving average components. In such cases, similarly to instances where negative R^2 values were encountered previously, employing the mean value proves to be a more reliable predictor.

An example of what has been just discussed is KPI3, but before analyzing its ARIMA model, it is recommend taking a step back and examine the ACF and PACF graphs, since this analysis could provide valuable insights into the underlying patterns and autocorrelations within the data.

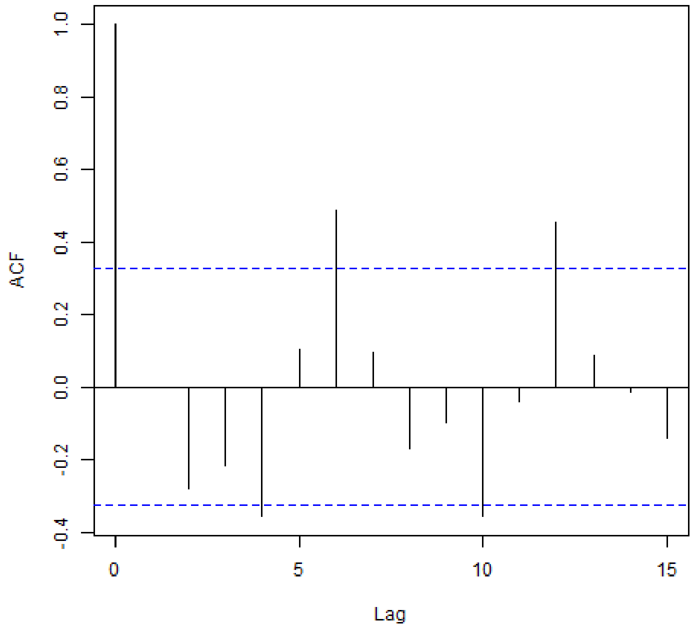

Figure 41: ACF function values by lags for KPI3

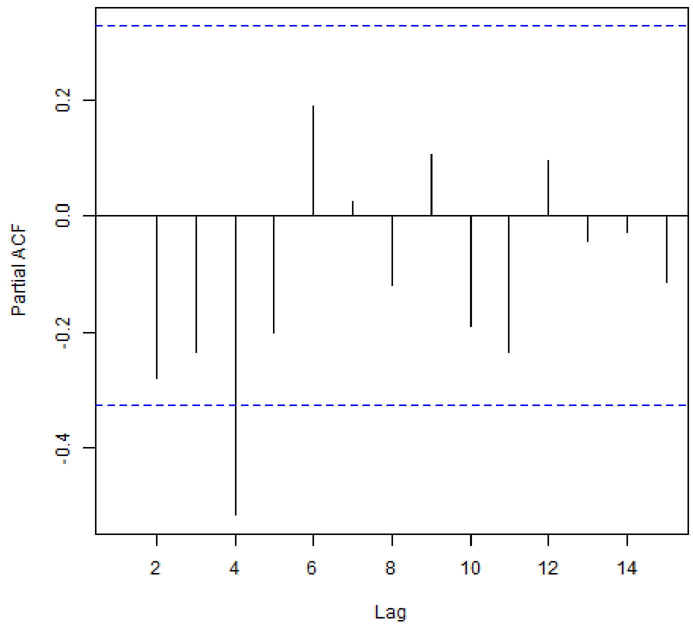

Figure 42: PACF function values by lags for KPI3

After having observed the two graphs above, it's evident that neither exhibits a distinct decay pattern, as underscored in chapter 2. Additionally, upon closer examination of the ACF plot, it is possible to notice discrepancies for certain lags, such as 5 or 13, where one graph registers a positive value while the other graph reports a negative one. These discrepancies could derive from various reasons, one of which may be the autoregressive model's poor fit with the given data. For instance, a significant negative value in the PACF implies that lag 4 remains unexplained by previous lags. Meanwhile, the ACF highlights five significant values, three positives and two negatives, suggesting correlations between current and lagged observations, some positive and others negative for different lags. In summary, while seasonality or other complex trends might explain the previous results, the time series appears void of any distinct patterns or trends that should require further exploration.

Now, moving our focus to the ARIMA graph:

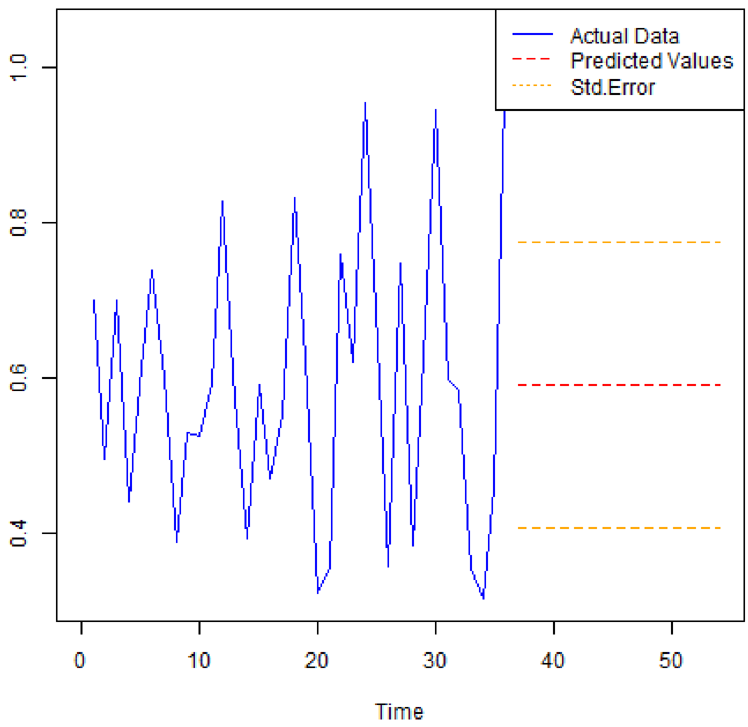

Figure 43: Time series graph and ARIMA prediction with standard error intervals for KPI3

With the previous insights it is possible to notice that the high variability, coupled with the limited data points, poses challenges for accurately predicting future data points, as evidenced by the wide standard error. For instance, under the assumption of normal distribution, one standard error from the mean should comprehend approximately 60% of all observations. However, as observed in KPI3, this is not the case, underlying the difficulty in predicting outcomes with certainty.

On the contrary, certain KPIs, like KPI2 and KPI11, reported all four factors, implying not only the presence of one coefficient for autoregressive and on for moving average components, but two for each. This suggests the existence of multiple autoregressive or moving average trends within the training data, and that are likely to persist into the future. This supposition finds support when examining their ACF and PACF plots. For example, let’s consider KPI11:

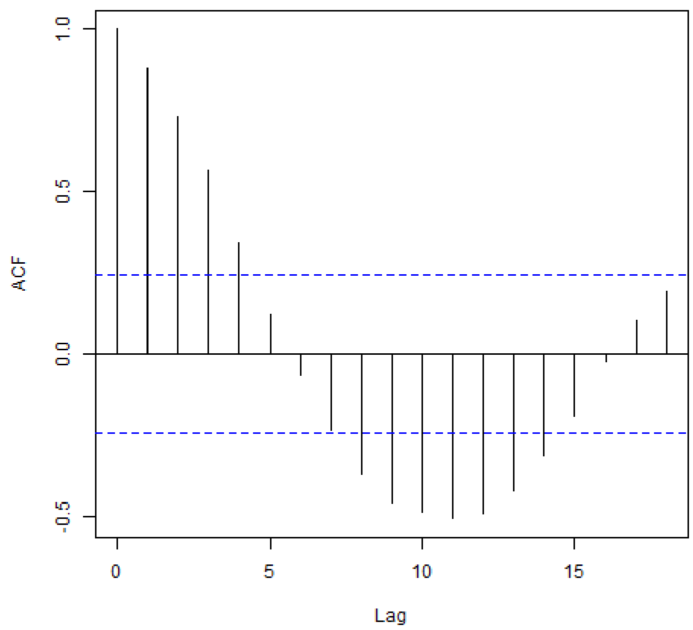

Figure 44: ACF function values by lags for KPI11

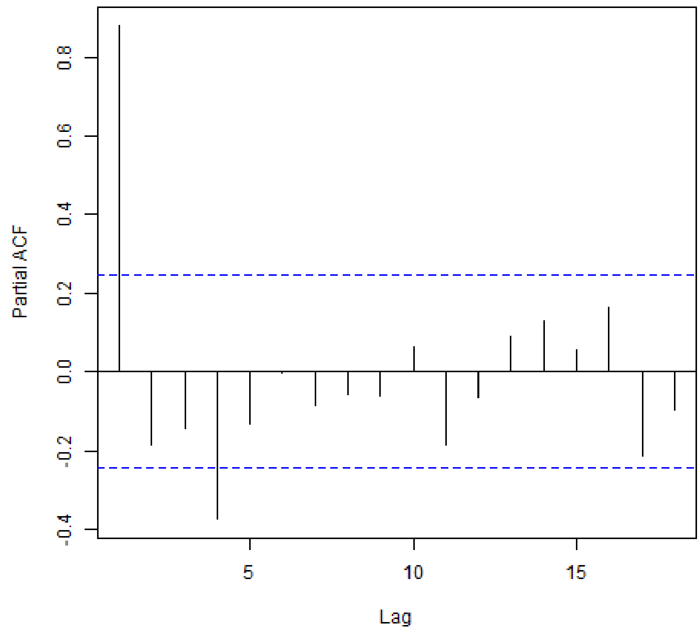

Figure 45: pACF function values by lags for KPI11

In contrast to KPI3, where trends are less defined, KPI11 reveals a clear cosinusoidal tendency in its ACF plot. This indicates potential seasonality in the data, corroborated by positive autocorrelation, AR(1) in the table, with the previous lag and negative correlation, AR(2), with observations two time points prior. Similarly, the negative correlation observed in MA(1) with the previous observation and positive correlation with the second previous observation, MA(2), support this conclusion.

However, despite the presence of numerous significance bars in the ACF plot, sustaining the non-random behavior hypothesis, their frequency is not as pronounced as in the PACF plot. This observation may clarify why the actual model reports a seasonality trend that diminishes over time, eventually converging towards the mean without reaching it. This phenomenon becomes apparent when inspecting the actual graph:

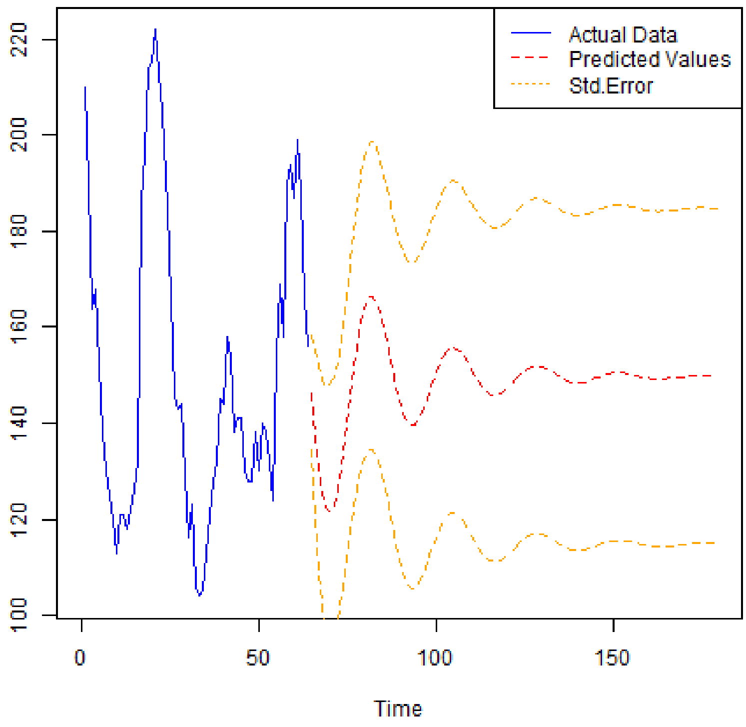

Figure 46: Time series graph and ARIMA prediction with standard error intervals for KPI11

The scenario for KPI2 presents some distinctive characteristics, particularly evident in its ACF and PACF graphs:

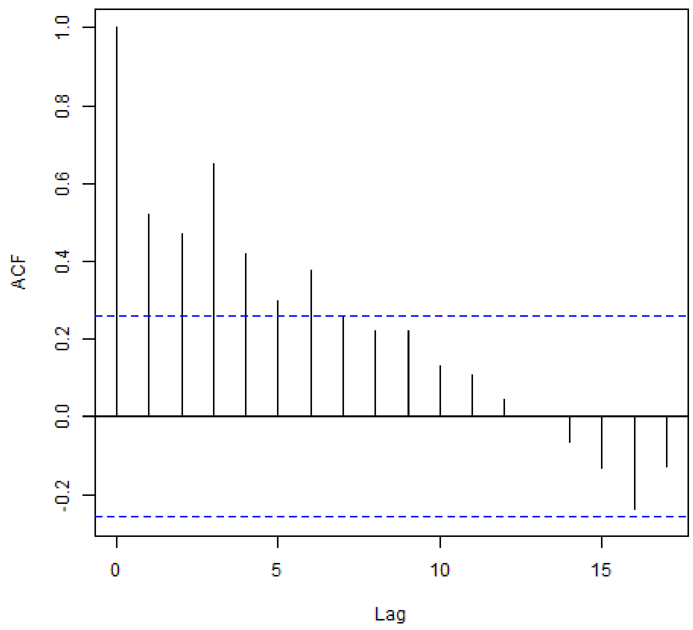

Figure 47: ACF function values by lags for KPI2

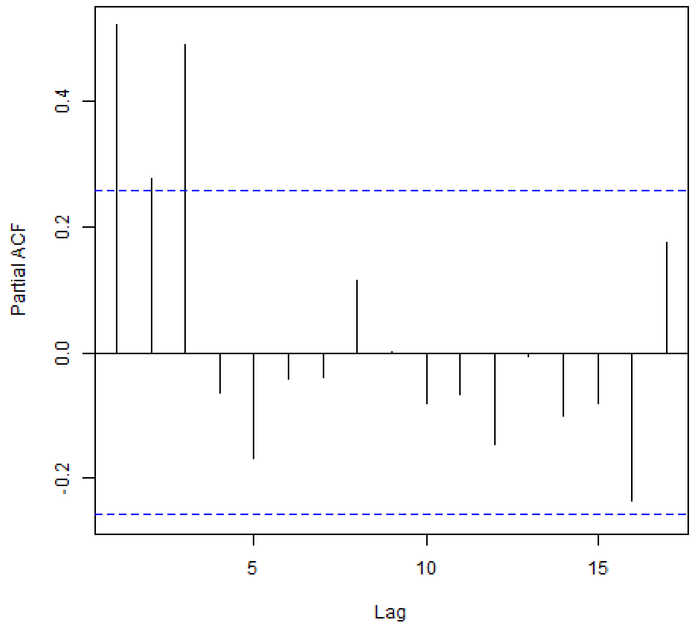

Figure 48: PACF function values by lags for KPI2

Upon examining these graphs, the notable feature is the negative slopes that seem to be followed by the bars in the ACF plot. Although a peculiar pattern emerges with every third lag in the ACF standing out from the subsequent two and breaking this monotone decreasement. This is an indicator or seasonality. However, the final lag in the ACF graph signals a deviation from this trend, hinting at a regression towards the mean, especially considering the diminishing positive correlations with increasing lags.

Conversely, the PACF plot does not show a similar trend of diminishing correlations with increasing lags, implying that the influence of each lag on the time series future tendency remains relatively stable once intermediate lags have been crossed out. However, it's worth noting that none of these lags reach significant values. This can be potentially attributed to chances.

To better visualize these dynamics, let's explore the ARIMA time series prediction for KPI2:

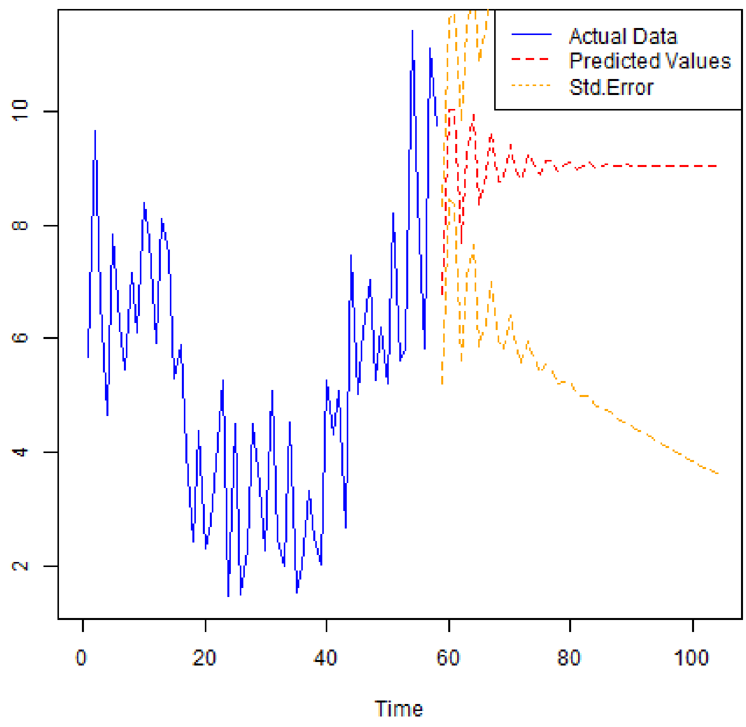

Figure 49: Time series graph and ARIMA prediction with standard error intervals for KPI2

Here, we can observe the ARIMA model's ability to identify patterns and trends, particularly in relation to the historical dataset, which is evident in KPI2's extensive number of previous observations. In particular, the variance in standard errors is strikingly big. While KPI11 exhibits a consistent pattern over time, the fluctuations in historical data preceding the project's commencement make it challenging to predict the trajectory of KPI2's performance long into the future.

Moreover, there are KPIs that exhibit both some autoregressive (AR) and moving average (MA) coefficients, along with a mean value component. This occurs when initially the KPI may display a general trend characterized by autoregressive or moving average patterns, but then fall precisely towards the mean and reach that. An illustration of this is KPI12, whose ACF and PACF graphs are presented below:

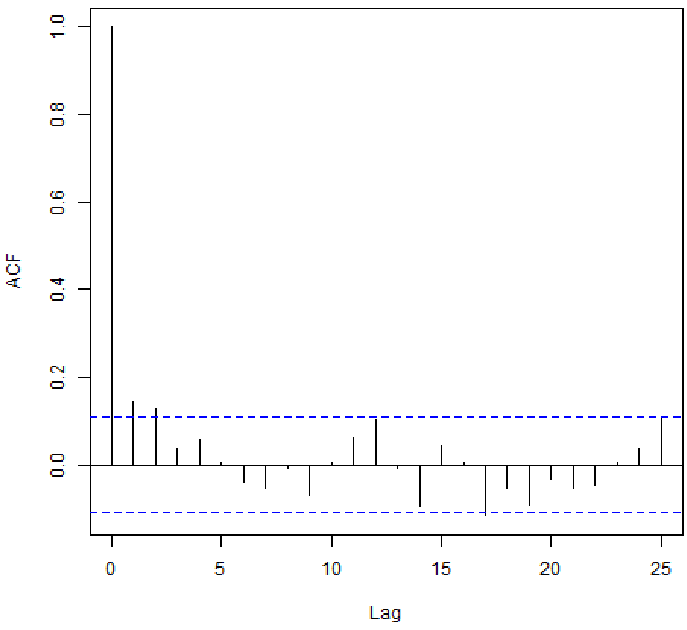

Figure 50: ACF function values by lags for KPI12

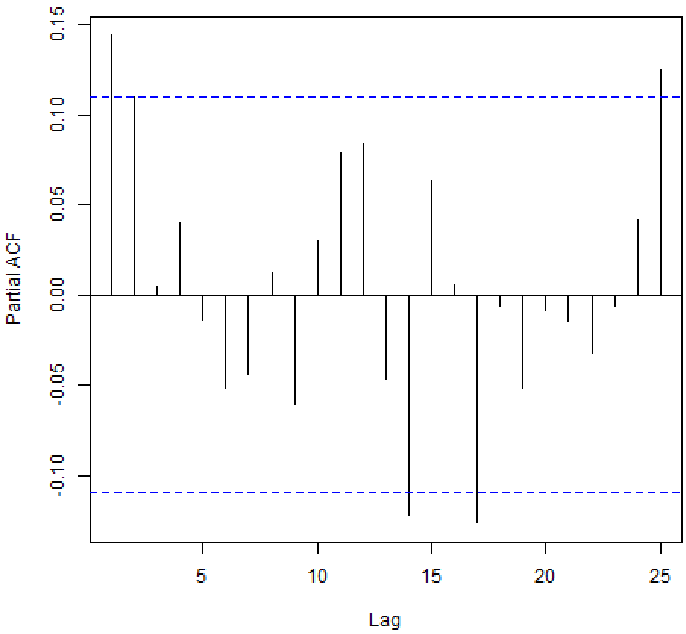

Figure 51: PACF function values by lags for KPI12

In the ACF plot, certain lags show small yet significant correlations, notably the first and second lags where the bars surpass the significance threshold. Regarding the PACF plot, several lags appear to contribute to the overall trend of the time series, such as lags 1, 2, 14, 17, and 25. However, there is no distinct decay or noteworthy pattern to report. So, we can deduct that a AR(s) trend can be expected while a MA is unlikely to verify.

Now, let's delve into the ARIMA prediction model:

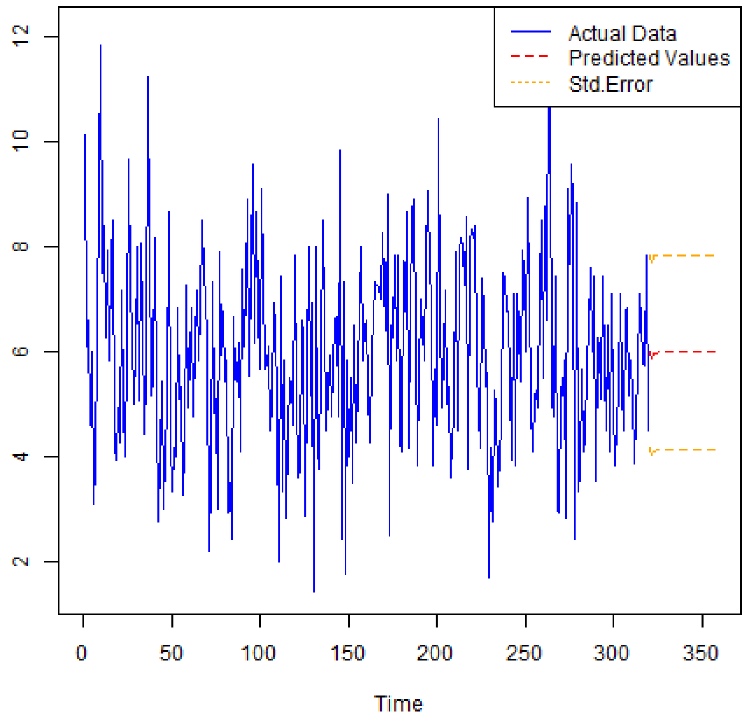

Figure 52: Time series graph and ARIMA prediction with standard error intervals for KPI12

As observable, there is a slight downward trend in the projected future values before reverting towards the mean. This suggests that an AR factor, in this case, influenced the prediction of the future trend without significantly altering the overall trend, which is characterized by small fluctuations around the mean, as indicated by the relatively low variance of the white noise.

From a managerial standpoint, time series analysis offers invaluable insights that extend beyond merely identifying trends or the anticipated direction of performance. It considers numerous factors crucial for assessing KPIs performances over time. Consequently, it's imperative for managers to analyze expected performances against actual values and adjust their current models accordingly.

Moreover, the significance of time series analysis in production cannot be overstated, as it accounts for various market or production-related factors that may cause sudden or temporary changes in production. Hence, by employing methods such as autoregressive and moving average techniques, managers can effectively address and considerate these conditions into their analyses.

This approach empowers managers to gain a deeper understanding of data and fosters clear, predictable performance in day-to-day activities. Such insights enable comparisons with clients and consultants, facilitating improvements in client performance and enhancing model predictability.

4.7. Consideration on risk related model

The subsequent step was to evaluate the viability of the risk-related model and construct a potential financial projects model. In order to do this, it was imperative to analyze the feasibility of the mean-variance model. This calls for a meticulous examination of some of the underpinning assumptions of the mean-variance model.

Primarily, the modern portfolio theory requires a multivariate normal distribution of returns. Hence, in our condition, we could verify it by assessing the performance distribution using the standardized Z-scores across all KPIs. However, as clarified in the standardization and Z-score section, the results of the analysis exhibit a left-skewed pattern, which is a significant challenge to the Markowitz model's foundational assumption. Because without the normal distribution the results of the model would not be representative of the actual reality.

Secondly, due to the nature of client data, the limitations evaluated in the section concerning financial statement analysis, and the absence of market analysis for the specific clients' sector, evaluating the outliers and extreme events would have been impractical in the best case, and impossible in the worst one. Without access to this data, any analysis must rely solely on assumptions, the validity of which depends entirely on the judgment of those making them. Unique businesses, target consumers, and limited public data reduce even the potential for conducting a proper market-wide analysis, raising other obstacles to execute an effective assessment.

Thirdly, projects tied to variable compensation, might present limited and dubious data points performances, as is the case for client number 4. In that case, the absence of historical data obstruct the possibility of running statistical testing, jeopardizing the model's robustness and making the eventual implementation of these statistical outcomes not advisable.

The solidity of historical data also introduces concerns regarding market dynamics and their influential impact on individual KPI performances. Seasonality, external conditions, and potential correlation between industries performances as a whole cannot be explored, because of this, so further issues arise concerning the reliability of the mean-variance model.

These are just a few potentials impairs of the mean-variance model, others are technological synergies, resource constraints, and dynamic risk parameters associated with specific projects. It follows that the assumption that risk and expected return are directly related may not hold true in some cases where, for instance, the contractual obligations limit adjustments of compensations if the risk varies.

Given these complexities, computing a mean-variance model appears unfeasible and lacks utility in the evaluation of potential variable compensation. Even in a simplified model considering only mean and variance, unreliable data, as evidenced in subsection 4.1.3, undermines the model's credibility across various categories. Developing such a model under these conditions would likely yield unsatisfactory and inconclusive results for fischer Consulting.

In light also of the complications identified in previous sections, it is advisable to hold back from pursuing the development of the mean-variance risk-related model, until these issues are addressed. However, if a broader set of observations was analyzed and the results proved compelling then the mean-variance model would be a valuable tool for fischer Consulting, enhancing their competitiveness in the market.

Conclusion

In conclusion, this thesis delves into the intricate realms of consultancy and statistics, employing robust statistical methods to evaluate the impacts that consultancy may have on clients' KPIs. It is imperative to note that the study derives its information from real-life data, sourced from companies that partnered with fischer Consulting Italia. The thesis not only serves as a nexus between statistics and consultancy but also scrutinizes the potential risks and complexities that real consulting firms must navigate when considering the introduction of a variable compensation system which is tied to performance related metrics.

Findings reveal that an underlying constraint in assessing the robustness of our results lies in the quantity and in the reliability of data available for analysis. For instance, the non-representativeness led us to the exclusion of KPI10, because of the excessive noise that it was introducing within the statistical analysis. Moreover, the statistical significance of our results varies notably across different KPIs and clients.

Initially, descriptive statistics were computed, yielding significantly positive results, evidenced by substantial changes in both weighted and unweighted target achievement. However, it is crucial to recall the magnitude of standard deviation on these conclusions. As illustrated in the attachments the TA’s improvements vary considerably across KPIs and categories, making the underlying conclusion impactful and questionable. It is worth remembering that the TA worsened in the productivity category when comparing the two periods (before and after). However, a remarkable improvement in efficiency was recorded, which was also the area with the highest number of observations. Additionally, while results for quality were positive, further investigation must be conducted to validate the solidity of the computed values, given the limited number of observations.

Another insightful analysis involved computing the Z-Score, which revealed a modest, yet significant, shift in its distribution towards normality. This underscores fischer Consulting's efficacy in enhancing predictability regarding business-related KPI performances, particularly in the analyzed sectors: efficiency, quality, and productivity.

The subsequent step involved evaluating the correlation between the recording dates and actual performance. After having recognized the absence of causation between these variables, this analysis served to statistically test the strength of our results, revealing, for the most part, no statistically significant correlation.

Our attention then shifted to examining instead the linear correlation between variables. One notable outcome was the considerable shift in result solidity between the before and after periods. Since these regressions were run separately, it was possible to observe numerous coefficients transitioned from insignificance to significance, highlighting the improvement of both the model and statistical analysis in explaining recorded results and performances. These results were further corroborated by the R^2 and adjusted R^2 values. These patterns were confirmed once again by categorical-level analysis, where a reversal of signs in adjusted R^2 was observed, transitioning from negative before to positive afterwards.

Nevertheless, it's worth noting the ambivalent results obtained during the more comprehensive analysis, which demonstrate statistical insignificance regarding certain variable correlations, namely *performance* and *date_obs*, while showing significant correlations with others, *performance* and *after*, interaction. This underscores the power of our comprehensive model, and control variables, which are then further explored through residual analysis. This analysis uncovered peculiarities, allowing us to assess the presence of heteroskedasticity, homoscedasticity, or neither of the two. A particular focus was dedicated to client four's KPIs and KPI11 residuals.

These insights were pivotal in driving more precise assessment, which raised questions about whether the linear model was indeed optimal for representing our data points. Consequently, a LOESS regression was conducted to explore potential non-linear relationships.

The non-parametric regression unveiled significant disparities between tendencies observed at the KPI and categorical levels. Specifically, the standard error was more contained in the KPI-level analysis, raising questions into result heterogeneity within the same category and its impact on possible macroscopic inferences. At a categorical level some bigger disparities have arisen, and some of these can be partly attributed to different time frames covered by each KPI's project, meaning the range of dates before and after project commencement. Nonetheless, intriguing non-linear trends were captured and explained concerning the case of KPI11.

Subsequently, an ARIMA model was employed to forecast future KPI specific performance trends. What was observed was that while useful predictions were generated for a limited number of KPIs, a substantial proportion of these lacked sufficient data for more accurate predictions beyond mean values. This validates the results obtained from previous analysis. However, with KPI that contains more observations, a deeper analysis revealed highly insightful predictions and cyclical insights, providing precise fitted values through reliable statistical tools. This analysis was enhanced by the utilization of ACF and PACF plots, which proved crucial in evaluating conclusive results.

The final consideration centered on the potential development of a variable fee model based on the Markowitz modern portfolio theory, which offers a promising prospect for future investigation and elaboration, as to reach an effective and advantageous model, but for now many complexities still need to be addressed.

From a statistical perspective, our analysis has yielded numerous key insights that could significantly improve fischer Consulting's approach to further analysis and risk assessment related with introducing variable components into their contracts. Leveraging the computational capabilities of R and meticulously analyzing the data have provided invaluable insights. One such discovery is the role of TA as a pivotal connector between multiple KPIs, making the comparison easier than what it would have been otherwise. In addition, we've gained valuable perspectives into the distribution of these KPIs' performances, enabling us to assess the probable outcomes and identify, even with some limitations due to variability considerations, the likelihood of achieving specific TA targets.

Of paramount importance is also our analysis's identification of areas of strength and areas requiring improvement within fischer Consulting's team. This knowledge empowers fischer's managers to locate additional resources to make up for the recorded gap, or organize targeted training programs, to address weaknesses effectively. Furthermore, it makes the entire team more aware of the inherent fluctuations in certain KPIs categories and raises questions about the potential impact of external factors such as market conditions and client-specific circumstances.

Another critical conclusion is the imperative for data collection, analysis, and curiosity. Through the use of the newly developed R software, fischer Consulting can easily integrate new data and receive continuous statistical evaluations. This iterative process amplifies the accuracy of our analysis' results and equips the team with the possibility to implement the mean-variance model effectively, allowing managers to make informed, data-driven decisions. This strategic approach not only could raise fischer Consulting's compensation, but also push for organic growth and improve its competitive edge in the consultancy sector. In particular, it positions fischer Consulting as a growing player with specialized expertise in manufacturing and process enhancement, leveraging the powerful tools of lean methodologies.

Making use of the transformative potential of statistical tools in guiding business decisions, challenging conventional wisdom in a country like Italy, which struggles with modernization and where innovation is often perceived as the domain of industry giants, will be a key factor in fischer Consulting's success and raise, and will demonstrate its efficacy as a dynamic player in this landscape.

However, it's essential to acknowledge that there are many areas that require further improvements. But we should not be concern about it, since any problem has its own solution and as the father of the TPS (Toyota Production System), Taiichi Ohno, once said:

"Having no problems is the biggest problem of all."

Acknowledgements

I extend my wholehearted gratitude to Prof. Maria Barbati for her invaluable guidance and unwavering support throughout the drafting of this thesis, as well as during my internship over the past months. Her dedication and time are, and have been, deeply appreciated.

I am intensely grateful to fischer Consulting Italia for providing access to the data that made this thesis possible, and to the entire team for the invaluable lessons imparted. My thanks extend to Federico Lazzarini, Dimitri Viale, Alessandro Mercuri, Giacomo Ambrosi, and Francesco Codognotto. A particular appreciation goes to Davide Bocca for his support and kindness throughout my internship, and the infectious enthusiasm with which he works.

A special mention goes to my dad, whose constant support has been an unshakeable source of strength, support, and motivation throughout the years. Much of what I've achieved is due to you. I am equally grateful to my grandfather, Raimondo, whose contagious laughter made me appreciate life more.

Last but certainly not least, I want to express my unparalleled appreciation to my best friends: Mattia, Pietro, Davide, and Giorgio. They have been my companions in countless adventures, offering support, advice, and sharing sincere joy in all my accomplishments.

To each of you and everyone else who has been close and supportive throughout these years, I am indebted. Without your support, encouragement, and friendship, I would not be the person I am today. Thank you, sincerely, from the deep of my mind.

Bibliography

Gaetano B. Ronsivalle. (April 2022). *La nuova intelligenza digitale. Come trasformare i dati in decisioni per progettare il futuro*, pp. 83-90.

Lyle D. Broemeling. (1st November 2011). "An Account of Early Statistical Inference in Arab Cryptology." *The American Statistician*, vol. 65, no. 4, 2011, pp. 255–257. doi:10.1198/tas.2011.10191. S2CID 123537702.

James P. Womack, Daniel I. Jones. (1st August 1996). *Lean Thinking: banish waste and create wealth in your corporation*.

Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying E. Ye. (2016). *Probability & Statistics for Engineers & Scientists*, 9th ed., pp. 27.

Rob J. Hyndman, Yeasmin Khandakar. (2008). "Automatic time series forecasting: The forecast package for R." *Journal of Statistical Software*, vol. 27, no. 1, 2008, pp. 1–22. doi: 10.18637/jss.v027.i03.

David A. Dickey, Wayne A. Fuller. (June 1979). "Distribution of the Estimators for Autoregressive Time Series With a Unit Root." *Journal of the American Statistical Association*, vol. 74, no. 366, 1979, pp. 427–431.

Denis Kwiatkowski, Peter C.B. Phillips, Peter Schmidt, Yongcheol Shin. (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of Econometrics*, vol. 54, no. 1-3, 1992, pp. 159-178.

Nick Huntington-Klein. (2021). *The Effect: An Introduction to Research Design and Causality*, ch. 4, sec. 4.4. doi: 10.1201/9781003226055.

João Paulo Figueira. (24th May 2019) "LOESS: Smoothing data using local regression." *Towards Data Science*.

Rob J. Hyndman, George Athanasopoulos. (April 2018). *Forecasting: Principles and Practice*, ch. 8, sec. 8.7.

George A. Akerlof. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, vol. 84, no. 3, 1970, pp. 488–500. doi:10.2307/1879431. JSTOR 1879431.

Walter A. Shewhart. (1986). *Statistical Method from the Viewpoint of Quality Control*.

David Simchi-Levi, Michael A. Trick. (1st June 2011). "Introduction to 'Little's Law as Viewed on Its 50th Anniversary'." *Operations Research*, vol. 59, no. 3, 2011, pp. 535. doi:10.1287/opre.1110.0941.

Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, Gert G. Wagner. "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey." *Journal of the European Economic Association*, vol. 9, no. 3, 2011, pp. 522–550.

Harry Markowitz. (1952). “Portfolio Selection.” *The Journal of Finance*, vol. 7, no. 1, 1952, pp. 77–91. JSTOR, doi: 10.2307/2975974.

Paul Schoemaker. (January 1995). "Scenario Planning: A Tool for Strategic Thinking." *Sloan Management Review*, vol. 36, no. 2, pp. 25-40.

Webography

Steve Skiena. (4th June 2001). "Probability versus Statistics, Calculated Bets: Computers, Gambling, and Mathematical Modeling to Win!" [Online]. Available: link.

Alicia Raeburn. (1st July 2023). "What's the difference between accuracy and precision?" [Online]. Available: link.

Quality-One. "Gage Repeatability & Reproducibility (Gage R&R)" [Online]. Available: link.

Simplilearn. "What is Descriptive Statistics: Definition, Types, Applications, and Examples" [Online]. Available: link.

Laerd Statistics. "Pearson Product-Moment Correlation" [Online]. Available: link.

Leonie Monigatti. (2nd August 2022). "Interpreting ACF and PACF Plots for Time Series Forecasting: How to determine the order of AR and MA models". *Towards Data Science*. [Online]. Available: link.

Viren Renhal. (21st June 2022). "Interpreting ACF and PACF plots". *Spur Economics*. [Online]. Available: link.

Eric. (23rd September 2020). "Beginner's Guide To Maximum Likelihood Estimation" [Online]. Available: link.

Gartner Inc. "Diffusion of Personal Computing Devices, 1977-2021." *The geography of transport system*. [Online]. Available: link.

R documentation

lm. stats. R Documentation. [Online]. Available: link.

pt. TDist, stats. R Documentation. [Online]. Available: link.

summary. base. R Documentation. [Online]. Available: link.

filter. dplyr. R Documentation. [Online]. Available: link.

abs. SparkR. R Documentation. [Online]. Available: link.

ts. stats. R Documentation. [Online]. Available: link.

cor.test. stats. R Documentation. [Online]. Available: link.

as.Date. base. R Documentation. [Online]. Available: link.

acf. stats. R Documentation. [Online]. Available: link.

pacf. stats. R Documentation. [Online]. Available: link.

auto.arima. forecast. R Documentation. [Online]. Available: link.

Appendix:

ATTACHMENT 1: Dataset variables with description and type of variable.

NAME	DESCRIPTION	TYPE
firm_name	Represents the firms' name	String
prj_code	Identification code of the particular project	String
date_start_prj	Date of beginning of the project.	Date
date_end_prj	Date of ending of the project.	Date
KPI_id	Identification code of a particular KPI	String
index_KPI	Identification of the specific observation for each KPI	String
date_obs	Date in which the observed value was recorded	Date
performance	Represents the recorded observed value	Float
target	Represents the aimed target	Float
KPI_des	Articulate description of the KPI	String
KPI_area	Identify the category of the tracked KPI	String
baseline	Represents the starting average value of the observed value of that peculiar KPI	Float
target_achievement	Represents the percentage of change with respect to the desired target change	Float

ATTACHMENT 2: Mean, median, and standard deviation of Performance and Target Achievement by period of observation (After and Before).

KPI id	After 1 = Yes 0 = No	PERFORMANCE			TARGET ACHIEVEMENT		
		Mean	Median	Std. Deviation	Mean	Median	Std. Deviation
KPI1	0	4.08989	3.94465	0.85767	0.00007	0.09142	0.53941
KPI1	1	4.36618	4.39974	0.52148	-0.17370	-0.19480	0.32797
KPI2	0	5.25129	5.25745	2.42821	-0.16638	-0.17102	1.82572
KPI2	1	6.26916	5.71625	2.70080	-0.93170	-0.51598	2.03068
KPI3	0	0.59099	0.59200	0.18508	0.00904	0.01814	1.68258
KPI3	1	0.59893	0.58728	0.11965	0.08118	-0.02470	1.08770
KPI4	1	0.88813	0.89000	0.04051	0.77762	0.81273	0.75854
KPI5	0	0.24352	0.26860	0.06210	0.07347	-0.17503	0.61525
KPI5	1	0.15434	0.15985	0.06491	0.95704	0.90241	0.64307
KPI6	0	0.19544	0.20090	0.06792	0.19640	0.09986	1.20106
KPI6	1	0.12364	0.11875	0.06270	1.46623	1.55264	1.10888
KPI7	0	5.97882	6.01000	0.41007	0.00120	-0.03061	0.41844
KPI7	1	5.57409	5.36000	0.83715	0.41419	0.63265	0.85424
KPI8	0	5.45059	5.35000	0.76664	-0.00131	0.22222	1.70364
KPI8	1	5.22409	5.29500	0.56278	0.50202	0.34444	1.25062
KPI9	0	23.79688	23.50000	6.65995	0.00427	-0.04338	1.06906
KPI9	1	21.73043	21.00000	6.17482	-0.32744	-0.44469	0.99119
KPI10	0	1.09375	1.00000	1.57075	0.13281	1.00000	14.52946
KPI10	1	0.85217	0.00000	1.31951	2.36739	10.25000	12.20549
KPI11	0	150.50000	141.00000	31.75701	0.20534	0.57360	1.23102
KPI11	1	135.34780	126.00000	43.40467	0.79270	1.15506	1.68253
KPI12	0	5.97740	5.84000	1.85287	0.01056	0.12605	1.55703
KPI12	1	5.17711	5.17000	2.01649	0.68310	0.68908	1.69452
KPI13	0	54314.60000	50683.00000	10193.88000	0.42661	0.51225	0.24039
KPI13	1	33941.29000	34557.00000	10482.16000	0.90706	0.89254	0.24719
KPI14	0	7.83951	2.73968	11.76740	-2.41976	0.13016	5.88370
KPI14	1	1.85788	1.54541	1.01590	0.57106	0.72729	0.50795
KPI15	0	3282905.00000	3333210.00000	227293.20000	0.02744	0.00000	0.12399
KPI15	1	2209779.00000	2149934.00000	907493.70000	0.61282	0.64547	0.49503

ATTACHMENT 3: Two tailed t-statistics by degrees of freedom and confidence interval.

	Area two-tailed										
df	0	0,5	0,6	0,7	0,8	0,9	0,95	0,98	0,99	0,998	0,999
1	0	1	1376	1963	3078	6314	12,71	31,82	63,66	318,31	636,62
2	0	0,816	1061	1386	1886	2,92	4303	6965	9925	22327	31599
3	0	0,765	0,978	1,25	1638	2353	3182	4541	5841	10215	12924
4	0	0,741	0,941	1,19	1533	2132	2776	3747	4604	7173	8,61
5	0	0,727	0,92	1156	1476	2015	2571	3365	4032	5893	6869
6	0	0,718	0,906	1134	1,44	1943	2447	3143	3707	5208	5959
7	0	0,711	0,896	1119	1415	1895	2365	2998	3499	4785	5408
8	0	0,706	0,889	1108	1397	1,86	2306	2896	3355	4501	5041
9	0	0,703	0,883	1,1	1383	1833	2262	2821	3,25	4297	4781
10	0	0,7	0,879	1093	1372	1812	2228	2764	3169	4144	4587
11	0	0,697	0,876	1088	1363	1796	2201	2718	3106	4025	4437
12	0	0,695	0,873	1083	1356	1782	2179	2681	3055	3,93	4318
13	0	0,694	0,87	1079	1,35	1771	2,16	2,65	3012	3852	4221
14	0	0,692	0,868	1076	1345	1761	2145	2624	2977	3787	4,14
15	0	0,691	0,866	1074	1341	1753	2131	2602	2947	3733	4073
16	0	0,69	0,865	1071	1337	1746	2,12	2583	2921	3686	4015
17	0	0,689	0,863	1069	1333	1,74	2,11	2567	2898	3646	3965
18	0	0,688	0,862	1067	1,33	1734	2101	2552	2878	3,61	3922
19	0	0,688	0,861	1066	1328	1729	2093	2539	2861	3579	3883
20	0	0,687	0,86	1064	1325	1725	2086	2528	2845	3552	3,85
21	0	0,686	0,859	1063	1323	1721	2,08	2518	2831	3527	3819
22	0	0,686	0,858	1061	1321	1717	2074	2508	2819	3505	3792
23	0	0,685	0,858	1,06	1319	1714	2069	2,5	2807	3485	3768
24	0	0,685	0,857	1059	1318	1711	2064	2492	2797	3467	3745
25	0	0,684	0,856	1058	1316	1708	2,06	2485	2787	3,45	3725
26	0	0,684	0,856	1058	1315	1706	2056	2479	2779	3435	3707
27	0	0,684	0,855	1057	1314	1703	2052	2473	2771	3421	3,69
28	0	0,683	0,855	1056	1313	1701	2048	2467	2763	3408	3674
29	0	0,683	0,854	1055	1311	1699	2045	2462	2756	3396	3659
30	0	0,683	0,854	1055	1,31	1697	2042	2457	2,75	3385	3646
40	0	0,681	0,851	1,05	1303	1684	2021	2423	2704	3307	3551
60	0	0,679	0,848	1045	1296	1671	2	2,39	2,66	3232	3,46
80	0	0,678	0,846	1043	1292	1664	1,99	2374	2639	3195	3416
100	0	0,677	0,845	1042	1,29	1,66	1984	2364	2626	3174	3,39
1000	0	0,675	0,842	1037	1282	1646	1962	2,33	2581	3098	3,3

ATTACHMENT 4: Summary results of complete regression of performance (dependent), with date_obs, after and date_obs:after interaction as regressors.

KPI id	Intercept	Coeff. Date of Observation	p-value Date of Observation	Coeff. of After	p-value Date of Observation	Coeff. of Interaction Date:After	p-value Date of Observation	R2	R2 Adjusted	AIC	BIC	RMSE
KPI1	20.3114	-0.0008	0.8183	2.6338	0.9816	-0.0001	0.9844	0.0383	-0.0441	94.1294	102.4472	0.7115
KPI2	-116.0410	0.0062	0.4028	862.9151	0.0006	-0.0441	0.0006	0.1651	0.1400	483.3144	496.5364	2.3552
KPI3	-1.6150	0.0001	0.8362	30.1070	0.3583	-0.0015	0.3585	0.0176	-0.0413	-33.3128	-23.3679	0.1620
KPI4	8.5887	-0.0004	0.9999	NA	NA	NA	NA	0.1282	0.1170	-291.9592	-284.8131	0.0376
KPI5	-3.0882	0.0002	0.6983	13.9410	0.1777	-0.0007	0.1764	0.4043	0.3533	-100.3316	-92.0138	0.0588
KPI6	14.9746	-0.0008	0.0997	-9.9247	0.3458	0.0005	0.3466	0.3116	0.2526	-98.4172	-90.0994	0.0603
KPI7	-41.7528	0.0025	0.5436	264.3275	0.0071	-0.0136	0.0073	0.4078	0.3570	72.1559	80.4737	0.5368
KPI8	31.1522	-0.0013	0.7648	114.7983	0.2648	-0.0059	0.2670	0.1856	0.1158	79.2349	87.5527	0.5878
KPI9	499.8641	-0.0245	0.3964	506.8134	0.4043	-0.0259	0.4080	0.1183	0.1032	1159.6177	1175.5547	6.0031
KPI11	-296.5973	0.0230	0.8947	8460.4495	0.0217	-0.4335	0.0222	0.1861	0.1722	1802.5767	1818.5136	36.1710
KPI12	7.5817	-0.0001	0.9071	459.2960	0.2201	-0.0236	0.2193	0.0214	0.0131	1470.7773	1490.1800	1.8613
KPI13	3935142.5931	-198.9862	0.7501	-1120306.1146	0.9275	57.0490	0.9280	0.6302	0.4915	260.8980	263.3225	8392.9318
KPI14	19541.0385	-1.0015	0.0184	-19282.9099	0.0207	0.9885	0.0207	0.5977	0.4468	81.2974	83.7219	4.7199
KPI15	-242415790.0730	12597.7796	0.7859	478892993.0463	0.6025	-24549.5894	0.6026	0.5595	0.3707	333.1481	335.1376	579320.3606