SOPRONI EGYETEM

LÁMFALUSSY SÁNDOR KÖZGAZDASÁGTUDOMÁNYI KAR

KONFERENCIAKÖTET

Conference Proceedings

Nemzetközi tudományos konferencia
a Magyar Tudomány Ünnepe alkalmából

International Scientific Conference
on the Occasion of the Hungarian Science Festival

Sopron, 2023. november 23.

23 November 2023, Sopron

FENNTARTHATÓSÁGI ÁTMENET: KIHÍVÁSOK ÉS INNOVATÍV MEGOLDÁSOK

SUSTAINABILITY TRANSITIONS: CHALLENGES AND INNOVATIVE SOLUTIONS

Szerkesztők / Editors:
OBÁDOVICS Csilla, RESPERGER Richárd, SZÉLES Zsuzsanna, TÓTH Balázs István

Nemzetközi tudományos konferencia a Magyar Tudomány Ünnepe alkalmából
International Scientific Conference on the Occasion of the Hungarian Science Festival

Sopron, 2023. november 23. / 23 November 2023, Sopron

FENNTARTHATÓSÁGI ÁTMENET: KIHÍVÁSOK ÉS INNOVATÍV MEGOLDÁSOK

SUSTAINABILITY TRANSITIONS: CHALLENGES AND INNOVATIVE SOLUTIONS

KONFERENCIAKÖTET

CONFERENCE PROCEEDINGS

LEKTORÁLT TANULMÁNYOK / PEER-REVIEWED PAPERS

Szerkesztők / Editors**:**
OBÁDOVICS Csilla – RESPERGER Richárd – SZÉLES Zsuzsanna – TÓTH Balázs István

SOPRONI EGYETEM KIADÓ

UNIVERSITY OF SOPRON PRESS

SOPRON, 2024

Nemzetközi tudományos konferencia a Magyar Tudomány Ünnepe alkalmából
International Scientific Conference on the Occasion of the Hungarian Science Festival

Sopron, 2023. november 23. / 23 November 2023, Sopron

A konferencia támogatói / Sponsors of the Conference:

Felelős kiadó / Executive Publisher**: Prof. Dr. FÁBIÁN Attila**
a Soproni Egyetem rektora / Rector of the University of Sopron

Szerkesztők / Editors:
Prof. Dr. OBÁDOVICS Csilla, Dr. RESPERGER Richárd,
Prof. Dr. SZÉLES Zsuzsanna, Dr. habil. TÓTH Balázs István

Lektorok / Reviewers:
Dr. habil. BARANYI Aranka, Prof. Dr. BÁRTFAI Zoltán, Dr. BARTÓK István, Dr. BEDNÁRIK Éva,
Bazsóné Dr. BERTALAN Laura, Dr. CZIRÁKI Gábor, Dr. DIÓSSI Katalin, Dr. habil. JANKÓ Ferenc,
Dr. KERESZTES Gábor, Dr. habil. KOLOSZÁR László, Dr. KÓPHÁZI Andrea,
Prof. Dr. KULCSÁR László, Dr. MÉSZÁROS Katalin, Dr. NEDELKA Erzsébet,
Dr. NÉMETH Nikoletta, Dr. NÉMETH Patrícia, Prof. Dr. OBÁDOVICS Csilla,
Dr. PALANCSA Attila, Dr. habil. PAPP-VÁRY Árpád Ferenc, Dr. RESPERGER Richárd,
Dr. habil. SZABÓ Zoltán, Prof. Dr. SZÉLES Zsuzsanna, Dr. SZÓKA Károly, Dr. TAKÁTS Alexandra,
Dr. habil. TÓTH Balázs István, Pappné Dr. VANCSÓ Judit

ISBN 978-963-334-499-6 (pdf)

DOI: 10.35511/978-963-334-499-6

SZERVEZŐK

Soproni Egyetem Lámfalussy Sándor Közgazdaságtudományi Kar (SOE LKK),
A Soproni Felsőoktatásért Alapítvány

A konferencia elnöke: Prof. Dr. SZÉLES Zsuzsanna egyetemi tanár, dékán (SOE LKK)

Tudományos Bizottság:

elnök:	Prof. Dr. OBÁDOVICS Csilla PhD egyetemi tanár, Doktori Iskola-vezető (SOE LKK)
társelnök:	Dr. habil. TÓTH Balázs István PhD egyetemi docens, igazgató (SOE LKK)
tagok:	Prof. Dr. FÁBIÁN Attila PhD egyetemi tanár (SOE LKK), rektor (SOE)
	Prof. Dr. SZÉKELY Csaba DSc professor emeritus (SOE LKK)
	Prof. Dr. KULCSÁR László CSc professor emeritus (SOE LKK)
	Prof. Dr. SZALAY László DSc egyetemi tanár (SOE LKK)
	Prof. Dr. Clemens JÄGER PhD egyetemi tanár, dékán (FOM)
	Dr. habil. BARANYI Aranka PhD egyetemi docens (SOE LKK)
	Dr. habil. POGÁTSA Zoltán PhD egyetemi docens (SOE LKK)
	Dr. habil. SZABÓ Zoltán PhD egyetemi docens (SOE LKK)
	Dr. habil. PAPP-VÁRY Árpád Ferenc PhD tudományos főmunkatárs (SOE LKK)
	Dr. Rudolf KUCHARČÍK PhD egyetemi docens, dékán (EUBA FIR)

Szervező Bizottság:

elnök:	Dr. RESPERGER Richárd PhD adjunktus (SOE LKK)
tagok:	Dr. KERESZTES Gábor PhD egyetemi docens, dékánhelyettes (SOE LKK)
	Dr. habil. Eva JANČÍKOVÁ PhD egyetemi docens (EUBA FIR)
	Dr. habil. KOLOSZÁR László PhD egyetemi docens, intézetigazgató (SOE LKK)
	Dr. HOSCHEK Mónika PhD egyetemi docens, intézetigazgató (SOE LKK)
	PAPPNÉ Dr. VANCSÓ Judit PhD egyetemi docens, intézetigazgató (SOE LKK)
	Dr. SZÓKA Károly PhD egyetemi docens (SOE LKK)
titkár:	NEMÉNY Dorka Virág kutatási asszisztens (SOE LKK)

ORGANIZERS

University of Sopron Alexandre Lamfalussy Faculty of Economics (SOE LKK),
For the Higher Education in Sopron Foundation

Conference Chairperson: Prof. Dr. Zsuzsanna SZÉLES PhD Professor, Dean (SOE LKK)

Scientific Committee:

Chair:	Prof. Dr. Csilla OBÁDOVICS PhD Professor, Head of Doctoral School (SOE LKK)
Co-Chair:	Dr. habil. Balázs István TÓTH PhD Associate Professor, Director (SOE LKK)
Members:	Prof. Dr. Attila FÁBIÁN PhD Professor (SOE LKK), Rector (SOE)
	Prof. Dr. Csaba SZÉKELY DSc Professor Emeritus (SOE LKK)
	Prof. Dr. László KULCSÁR CSc Professor Emeritus (SOE LKK)
	Prof. Dr. László SZALAY DSc Professor (SOE LKK)
	Prof. Dr. Clemens JÄGER PhD Professor, Dean (FOM)
	Dr. habil. Aranka BARANYI PhD Associate Professor (SOE LKK)
	Dr. habil. Zoltán POGÁTSA PhD Associate Professor (SOE LKK)
	Dr. habil. Zoltán SZABÓ PhD Associate Professor (SOE LKK)
	Dr. habil. Árpád Ferenc PAPP-VÁRY PhD Senior Research Fellow (SOE LKK)
	Dr. Rudolf KUCHARČÍK PhD Associate Professor, Dean (EUBA FIR)

Organizing Committee:

Chair:	Dr. Richárd RESPERGER PhD Assistant Professor (SOE LKK)
Members:	Dr. Gábor KERESZTES PhD Associate Professor, Vice Dean (SOE LKK)
	Dr. habil. Eva JANČÍKOVÁ PhD Associate Professor (EUBA FIR)
	Dr. habil. László KOLOSZÁR PhD Associate Professor, Director of Institute (SOE LKK)
	Dr. Mónika HOSCHEK PhD Associate Professor, Director of Institute (SOE LKK)
	Dr. Judit PAPPNÉ VANCSÓ PhD Associate Professor, Director of Institute (SOE LKK)
	Dr. Károly SZÓKA PhD Associate Professor (SOE LKK)
Secretary:	Dorka Virág NEMÉNY Research Assistant (SOE LKK)

TARTALOMJEGYZÉK / CONTENTS

3. szekció: Fenntartható pénzügyek és számvitel
Session 3: Sustainable Finance and Accounting

4. szekció: Fenntartható turizmus és marketing
Session 4: Sustainable Tourism and Marketing

5. szekció: Sustainable Finance and Accounting, Sustainable Development
Session 5: Sustainable Finance and Accounting, Sustainable Development

6. szekció: Sustainability Transformation and Circular Economy
Session 6: Sustainability Transformation and Circular Economy

7. szekció: Sustainable Economy and Management
Session 7: Sustainable Economy and Management

8. szekció: Társadalmi kihívások és társadalmi innovációk a fenntartható fejlődésben
Session 8: Social Challenges and Innovations in Sustainable Development

9. szekció: Fenntartható gazdálkodás és menedzsment
Session 9: Sustainable Economy and Management

10. szekció: Sustainable Economy and Management I.
Session 10: Sustainable Economy and Management I.

11. szekció: Sustainable Economy and Management II.
Session 11: Sustainable Economy and Management II.

12. szekció: Poszter szekció
Session 12: Poster Session

DOI: 10.35511/978-963-334-499-6-Toth-Garab

Examining the Impact of Certain Factors on the Delivery Time of a Manufacturing Firm Using Data Science Methods

Dr. Zsolt TÓTH PhD
Associate Professor
University of Sopron, Faculty of Wood Engineering and Creative Industries
toth.zsolt@uni-sopron.hu *(Corresponding Author)*

Dr. József GARAB PhD
Associate Professor
University of Sopron, Faculty of Wood Engineering and Creative Industries
garab.jozsef@uni-sopron.hu

Abstract
The study analysed a business big data problem in a case study and partly methodological way. In particular, the factors that influenced the order fulfilment time were investigated. From a table of nine columns and 15,388,470 rows, a time series analysis was first performed and then a Spearman correlation was used to analyse the relationship between completion time and the number of sales. The limitations of some methods (ANOVA, Random Forest, Factor Analysis of Mixed Data) were examined, and then Kruskal-Wallis and Dunn-Bonferroni tests were carried out, taking samples of sufficient and necessary size to allow for the large number of element variations in each column, and restricting the analysis to the variable values that sometimes appeared important. A number of variables in three columns had a significant effect on the evolution of the completion time and the methodology proved successful in identifying them. However, the analysis also highlighted the need to group the individual, non-metric variable values into new groups.

Keywords: big data, times series, completion time, mixed data
JEL Codes: C12, C55, L62

1. Introduction, objectives

The research was conducted with the support of an industrial partner, a mid-sized manufacturing enterprise. The study analysed a business big data problem in a case study and partly methodological way. In particular, the factors that influenced the order fulfilment time were investigated. The paper mainly provides a case study example of how big data problems can be solved with data science tools with limited resources.

Depending on the industry, the company, the production structure, the delivery time can be influenced by many factors. Due to the analytical constraints generated mainly by the large amount of data, as will be discussed later, really high-quality analyses have only been produced mainly in the last years.

For industrial companies producing relatively small batches, in addition to the production characteristics of the companies, the expected timing of order batches is the most important factor, and machine learning methods can achieve very significant results (Rokoss et al., 2024). A broader industry analysis also shows that collecting and analysing data more widely than before is essential to optimise delivery times. In addition to data on technical characteristics, one of the biggest threats is the failure to take into account human factors, in particular the needs of employees (Munmun et al., 2023). The efficient and general use of “historical” data to optimise and predict delivery times has led to the development of specific standardisation tools, such as Cross-Industry Standard Process for Data Mining (CRISP-DM) (Schuh et al., 2019). Optimizing the accommodations for variations in demand is an added benefit of a standardized implementation procedure for engineering improvements (Knaus, 2022). In addition to production systems (or data of these systems), the analysis and optimisation of related logistics systems can be particularly important in the management of large fluctuations (Härtel & Nyhuis, 2019). Supply chain management practices can have significant effects on delivery time (Anatan, 2014). In general terms, the uncertainty of delivery time can be understood primarily in terms of information uncertainty, not in technical terms (Busert & Fay, 2018).

2. Applied method and database

The initial data file was converted into a data frame R object. After data cleaning and recoding, we obtained a table with nine columns and 15.388.470 rows. Due to the large amount of data and the variety of categorical variables, the methods that can be used are mainly determined by the scarcity of computational capacity.

In the time series analysis, we first calculated the monthly average of completion times and the number of sales (sales transactions) per month, and then factorised the time series using an additive model. In our case, the delivery time is an external factor, as customers indicate in advance the delivery time they will order the product (Ulrich, 2021).

Nevertheless, it may be interesting to explore what factors - technological, economic or psychological - may influence this value. It is also worth comparing the components of the two time series. Since the completion time is metric as a dependent variable and the independent variables "customer", "article" and "auftrag" are non-metric, analysis of variance (ANOVA) seems to be an appropriate analytical tool. Due to the large number of data, a random sample of n = 100,000 was taken from the data. From the point of view of the ANOVA method, this does not imply a significant loss of data and the sampling error is negligible (Wang et al., 2020). In all three cases, for group *i* (*i* group/category is not the same for the three independent variables):

H_0: $\mu_1 = \mu_2 = \cdots = \mu_i$
H_1: At least two means differ.
Where μ_i is the average completion time.

It is worth noting that, although in practice data analysts tend to disregard this because of the large number of items, an important prerequisite for ANOVA is normality (Lantz, 2013), which is not fulfilled in our case. An alternative is to use the Kruskal-Wallis test (McKight & Najab, 2010). The Kruskal-Wallis test is based on analysis of variance in the same way as ANOVA, but tests the equality/differences of the median rather than the means through their ranks. In all three cases, for group *i*:

H_0: $Me_1 = Me_2 = \cdots = Me_i$
H_1: At least two medians differ.
Where Me_i is the median of completion time (Bolar, 2019).

To explore deeper correlations, the use of Random Forest (Schonlau & Zou, 2020) or FAMD (Factor Analysis of Mixed Data) (Bai et al., 2020) may be considered. However, these models would require a very large amount of computing power with such a large amount of data (Costa et al., 2023). Their applicability would be greatly increased if the large number of categories in the columns "customer", "article", "auftrag" (303, 1871 and 1535 items respectively) could be replaced by a small number of consolidated categories. In this case, the computational effort would be radically reduced, and any necessary sampling would not introduce any significant bias.

Pairwise comparison methods require less machinery (Kocak et al., 2018), but with such a large data set and many categories, most of them are either not computable with the available tools or can be used with a smaller number of categories.

The previous sampling procedure cannot be relied upon either, as some categories with a small number of items would probably be dropped from the sample, and if they were included with a single row, the sampling error would be very high (Alkhlaifat & Koloszár, 2023).

Therefore, we chose the method of selecting the ten most frequently occurring variables for each categorical variable, creating three filtered files, and then creating three n = 10000 samples from these. In this case, the sampling error is unlikely to be high in any of the categories, although the resulting samples are only suitable for pairwise comparisons of the ten most frequent categories for the ten most frequent categories per non-metric variable.

We also performed the Kruskal-Wallis test on the three samples. It can be clearly seen that, even for the categories with the ten to ten most rows in the samples, the effect of each category on the evolution of the completion time is true.

The specific effect is tested using the Dunn-Bonferroni test (Bonferroni correction), which is well-suited for both parametric and non-parametric procedures (Castañeda et al., 1993), which are essentially multiple Mann-Whitney tests (general H_0: the two groups under test are identical), where the significance level is divided by the *number of categories/groups - 1* (number of pairwise comparisons, or df). Thus α = 0.0056 (Ogle et al., 2023).

3. Results

The average completion time and the number of sales per month show a very similar pattern. The development of the steady and then the rising phases is similar (Figure 1).

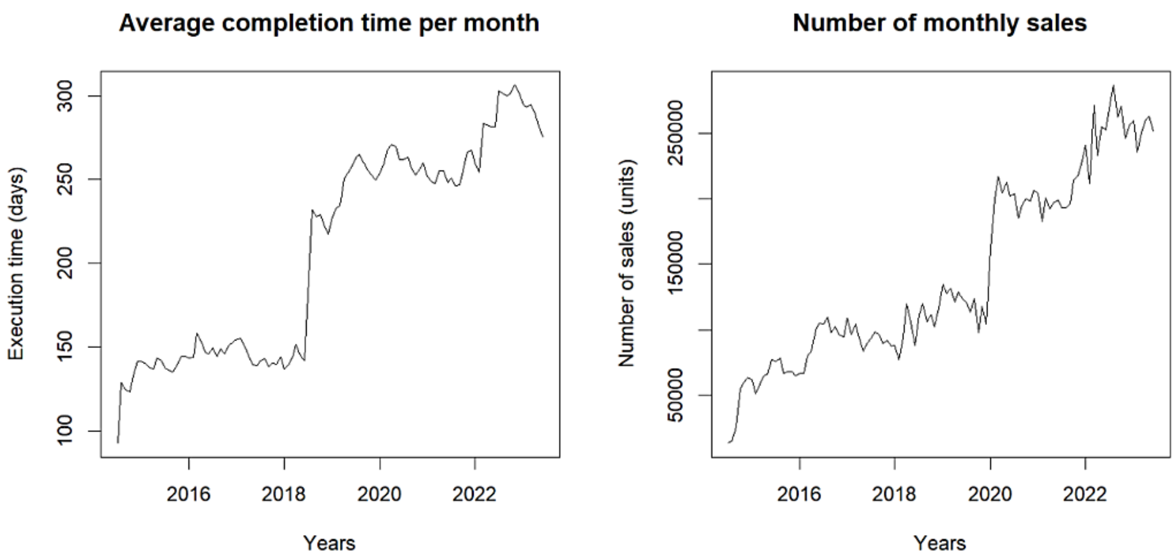

Figure 1. Original time series
Source: Own calculation (2023)

It is clear that the monthly completion time is increasing in a trend but not steadily. There is a strong negative seasonal effect in June and a strong positive seasonal effect in August. The negative peak in June 2018 and the positive peak in August 2018 appear to be outliers in the coincident effect on the monthly average completion time (which is obviously caused by some one-off effect). It is conceivable that the seasonal effect, which is present every year, was much stronger in this year, but the difference in magnitude does not suggest this (Figure 2).

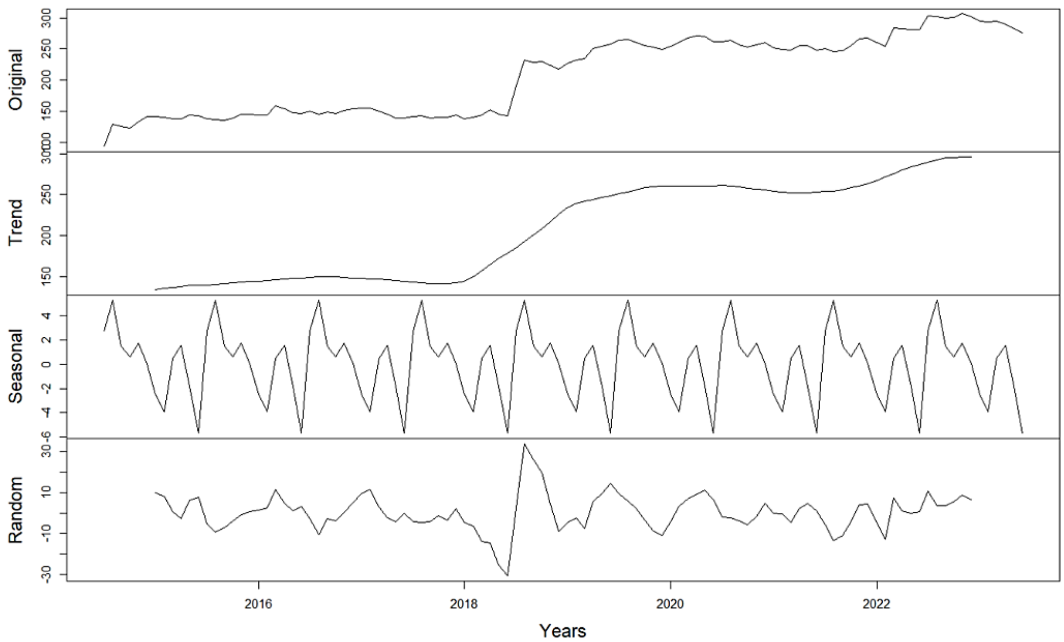

Figure 2. Monthly average completion time
Source: Own calculation (2023)

The trend in monthly sales is very similar to the average monthly completion time. However, the trend increase is more uniform. The seasonal effect shows a positive peak in March and a negative peak in November. A strong random effect is measured in February 2020 (positive direction) and December 2020 (negative direction) (Figure 3).

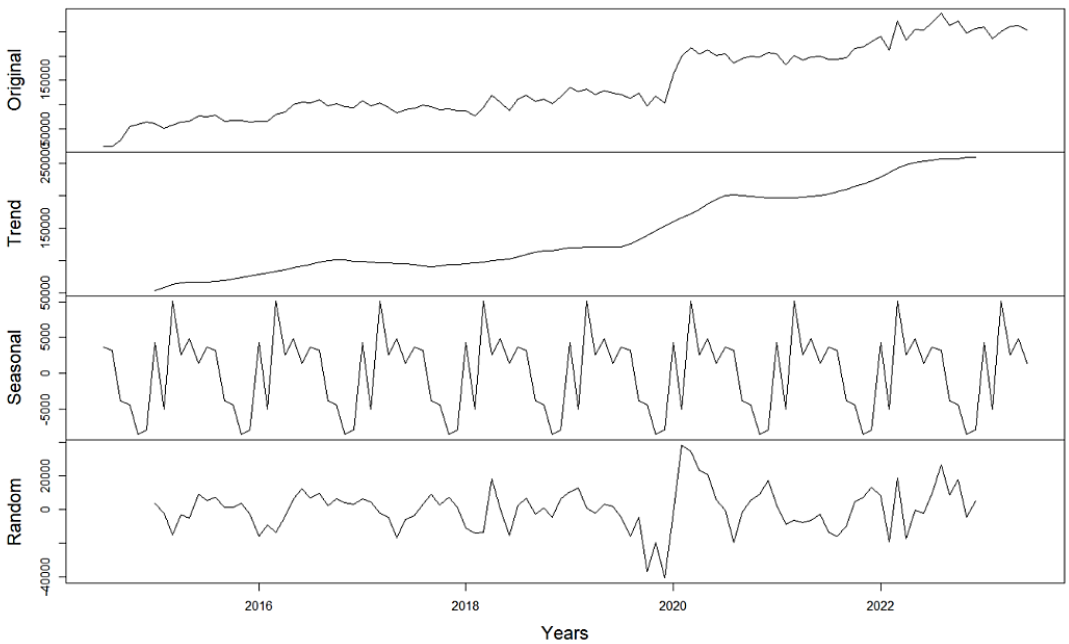

Figure 3. Number of sales

Source: Own calculation (2023)

The similarity of the two time series makes it worth comparing the initial monthly values. Due to the lack of normality, a non-parametric Spearman correlation can be calculated. The very high value of the correlation coefficient (ρ = 0.93) indicates a very close relationship between the two data series. When examining monthly sales, the quantity sold was not taken into account and only the number of transactions was calculated because the Spearman rank correlation between the time of completion and the quantity sold is -0.26, i.e. a weak to medium negative relationship.

Since $Pr(>F)$ (p) is less than $\alpha = 0.05$ in all three cases, the result is significant, i.e. there are significantly different means in each category, i.e. the completion time is influenced by the values of the three variables under investigation. However, this test can only detect the fact of the effect, not the extent (Table 1).

Table 1: ANOVA for n = 100,000 random samples

Ind. variable	df	Square sum	Average of sq. sum	F	Pr(>F) (p)
"customer"	200	8.165×10^{8}	4 082 379	179.5	$< 2 \times 10^{-16}$
residuals	99 799	2.270×10^{9}	22 744		
"article"	1 202	7.688×10^{8}	639 624	27.27	$< 2 \times 10^{-16}$
residuals	98 797	2.317×10^{9}	23 457		
"auftrag"	819	1.019×10^{9}	1 244 689	59.73	$< 2 \times 10^{-16}$
residuals	99 180	2.067×10^{9}	20 840		

Source: Own calculation (2023)

Due to the low p value, we can conclude that the three independent variables have an impact on the evolution of the completion time. This is as much information as can be obtained from this test, and further investigation is needed (Table 2).

Table 2: Kruskal-Wallis test for n = 100,000 random samples

Ind. variable	χ^2	df	p
"customer"	4 212 404	302	$< 2 \times 10^{-16}$
"article"	3 885 191	1 870	$< 2 \times 10^{-16}$
"auftrag"	5 057 068	1 534	$< 2 \times 10^{-16}$

Source: Own calculation (2023)

Therefore, we chose the method of selecting the ten most frequently occurring variables for each categorical variable, creating three filtered files, and then creating three n = 10000 samples from these.

Table 3: The ten categories with the most items for the three independent variables

Var. categories	Var. values	n (pop.)	n (sample)
"customer"	3405005	2 545 308	2 287
	3404999	2 340 213	2 058
	3403784	1 409 999	1 214
	3405097	1 370 881	1 220
	3405487	1 062 614	984
	3400154	646 991	539
	3403547	583 776	497
	3430705	547 642	461
	3403139	486 227	440
	3430938	322 324	300
"article"	1075881	184 621	1 169
	1074495	173 419	1 052
	1075856	171 681	1 093
	1076384	167 509	984
	1075431	164 550	992
	1074087	159 903	974
	1076073	157 682	999
	1075258	154 640	936
	1074077	152 679	871
	1076880	149 774	930
"auftrag"	101471	2 537 089	2 723
	101472	1 435 720	1 551
	3	1 062 480	1 116
	101182	869 724	962
	101530	803 365	924
	101426	555 471	588
	10	554 452	611
	101624	499 501	528
	101657	484 207	525
	101220	462 411	472

Source: Own calculation (2023)

In this case, the sampling error is unlikely to be high in any of the categories, although the resulting samples are only suitable for pairwise comparisons of the ten most frequent categories for the ten most frequent categories per non-metric variable. The element numbers of the samples generated by this method are shown in Table 3. It can be seen that the randomly sampled subsamples all have a rather high number of elements. Thus, the results of the statistical tests performed will be reliable in each sub-sample.

Table 4. Kruskal-Wallis test on a sample of n = 10,000 for the ten most populous categories

Ind. variable	χ^2	df	p
”customer”	2 473	9	$< 2x10^{-16}$
”article”	438	9	$< 2x10^{-16}$
”auftrag”	3 008	9	$< 2x10^{-16}$

Source: own calculation (2023)

The Kruskal-Wallis test clearly shows that even for the categories with the ten to ten most rows in the sample, the effect of each category on the evolution of the completion time is true. (Table 4) However, this does not mean that fewer variables and less data carry the same amount of information, but we can stochastically claim that the basic context has not changed drastically with the reduction. Of course, such statements should be treated with caution and only in the case of business and not primarily scientific analyses, such a methodological simplification can be accepted because of the technical bottlenecks.

The $p < \alpha$ values calculated from the Dunn-Bonferroni test suggest that the two variables have different effects on the completion time in a given pair. As a sample, pairs of ”customer” in Table 5.

Table 5. P values of the Dunn-Bonferroni test (completion time - "customer")

”customer”	3400154	3403139	3403547	3403784	3404999	3405005	3405097	3405487	3430705
3403139	0	-	-	-	-	-	-	-	-
3403547	≈0	≈0	-	-	-	-	-	-	-
3403784	≈0	≈0	≈0	-	-	-	-	-	-
3404999	≈0	≈0	0.63	≈0	-	-	-	-	-
3405005	≈0	1	≈0	≈0	≈0	-	-	-	-
3405097	≈0	≈0	1	≈0	0.26	≈0	-	-	-
3405487	1	≈0	≈0	≈0	≈0	≈0	≈0	-	-
3430705	1	≈0	≈0	≈0	≈0	≈0	≈0	1	-
3430938	≈0	1	≈0	≈0	≈0	1	≈0	≈0	≈0

Source: Own calculation (2023)

The above test, unlike the previous ones, already concretizes the relationships of the model with reduced variables and data. Of course, it is not the correlations between the specific variable values that are of interest, but the methodology, which is satisfactory even with the technical limitations, as it allowed us to find correlations and quantify them to some extent.

4. Conclusions

A number of variables in three columns had a significant effect on the evolution of the completion time and the methodology proved successful in identifying them. However, the analysis also highlighted the need to group the individual, non-metric variable values into new groups.

Without this change, only a very strong data reduction, after resampling, could lead to acceptable results compared to the analysis of the original data. It is worth noting that data analysis methods considered to be more advanced (e.g. Random Forest) also mostly do not consider all possible relations and simplify at other levels.

Our methodology, complemented by qualitative analysis, can reveal new knowledge within an organisation even when a supercomputer is not available. (In business, it is very rare for a company to use a supercomputer on a daily basis.)

The main focus is on qualitative analysis. If you do not carry out other qualitative analyses, but specifically reorganise individual variables into strict non-metric, i.e. qualitative categories, based on the internal functioning of the company, you can obtain results that are easier to interpret.

We could not make this change for the company. But we have tried to achieve a relatively acceptable result using data reduction methods. Both the method we used and the change we proposed go against the grain of current data science methods, but in our view this is the way forward for most companies.

The frantic pace of scientific expectations is very evident in data science analysis. New methods of analysis and new tests are expected of researchers. Methods become so often cited that no one uses them in practice. Meanwhile, the scientific value of old, tried and tested methods that are still in use today is unacceptably low. However, either the Kruskal-Wallis test or the Dunn-Bonferroni test are widely analysed, relatively 'old' but well-established methods. The aim of our study was to show an example of this.

References

Anatan, L. (2014). Factors Influencing Supply Chain Competitive Advantage and Performance. *International Journal of Business and Information*, *9*(3), 311–334.

Bai, H., Xie, N., Di, X., & Ye, Q. (2020). FAMD: A Fast Multifeature Android Malware Detection Framework, Design, and Implementation. *IEEE Access*, *8*, 194729–194740. https://doi.org/10.1109/ACCESS.2020.3033026

Bolar, K. (2019). *STAT: Interactive Document for Working with Basic Statistical Analysis* (0.1.0) [Computer software]. Retrieved: 2023.11.03, from https://cran.r-project.org/web/packages/STAT/index.html

Busert, T., & Fay, A. (2018). Information Quality Dimensions for Identifying and Handling Inaccuracy and Uncertainty in Production Planning and Control, In *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)* (pp. 581–588). https://doi.org/10.1109/ETFA.2018.8502465

Castañeda, M. B., Levin, J. R., & Dunham, R. B. (1993). Using planned comparisons in management research: A case for the bonferroni procedure. *Journal of Management*, *19*(3), 707–724. https://doi.org/10.1016/0149-2063(93)90012-C

Costa, E., Papatsouma, I., & Markos, A. (2023). Benchmarking distance-based partitioning methods for mixed-type data. *Advances in Data Analysis and Classification*, *17*(3), 701–724. https://doi.org/10.1007/s11634-022-00521-7

Härtel, L., & Nyhuis, P. (2019). Systematic Data Analysis in Production Controlling Systems to Increase Logistics Performance. In Schmitt, R., & Schuh, G. (Eds.), *Advances in Production Research* (pp. 3–13). Springer International Publishing. https://doi.org/10.1007/978-3-030-03451-1_1

Knaus, A. (2022). SECPON: Analysis of Engineering Change Processes for Manufacturing Companies. *International Journal of Innovation and Technology Management*, *19*(07), 2250027. https://doi.org/10.1142/S0219877022500274

Kocak, H., Caglar, A., & Oztas, G. Z. (2018). Euclidean Best–Worst Method and Its Application. *International Journal of Information Technology & Decision Making*, *17*(05), 1587–1605. https://doi.org/10.1142/S0219622018500323

Lantz, B. (2013). The impact of sample non-normality on ANOVA and alternative methods. *British Journal of Mathematical and Statistical Psychology*, *66*(2), 224–244. https://doi.org/10.1111/j.2044-8317.2012.02047.x

McKight, P. E., & Najab, J. (2010). Kruskal-Wallis Test. In Weiner, I. B., & Craighead, W. E. (Eds.), *The Corsini Encyclopedia of Psychology* (pp. 1–10). John Wiley & Sons. https://doi.org/10.1002/9780470479216.corpsy0491

Munmun, S. A., Moktadir, Md. A., Tiwari, S., Abedin, M. Z., & Jabbour, C. J. C. (2023). Investigation of key performance indicators for performance management of the manufacturing industry in the era of the COVID-19 pandemic. *Annals of Operations Research*. https://doi.org/10.1007/s10479-023-05717-4

Ogle, D. H., Doll, J. C., Wheeler, A. P., & dunnTest()), A. D. (Provided base functionality of. (2023). *FSA: Simple Fisheries Stock Assessment Methods* (0.9.5) [Computer software]. Retrieved: 2023.11.13, from https://cran.r-project.org/web/packages/FSA/index.html

Rokoss, A., Syberg, M., Tomidei, L., Hülsing, C., Deuse, J., & Schmidt, M. (2024). Case study on delivery time determination using a machine learning approach in small batch production companies. *Journal of Intelligent Manufacturing*. https://doi.org/10.1007/s10845-023-02290-2

Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. *The Stata Journal*, *20*(1), 3–29. https://doi.org/10.1177/1536867X20909688

Schuh, G., Prote, J.-P., Molitor, M., Sauermann, F., & Schmitz, S. (2019). Databased learning of influencing factors in order specific transition times. *Procedia Manufacturing*, *31*, 356–362. https://doi.org/10.1016/j.promfg.2019.03.056

Ulrich, J. (2021). *TTR: Technical Trading Rules* (0.24.3) [Computer software]. Retrieved: 2023.11.13, from https://cran.r-project.org/web/packages/TTR/index.html

Wang, F., Huang, G. H., Fan, Y., & Li, Y. P. (2020). Robust Subsampling ANOVA Methods for Sensitivity Analysis of Water Resource and Environmental Models. *Water Resources Management*, *34*(10), 3199–3217. https://doi.org/10.1007/s11269-020-02608-2