Data Measures Selection and Factor Profiling:

An Illustration of Data-Mining Screens

Mohamed Gaber[1] & Edward J. Lusk[2]

[1]Department of Accounting, School of Business and Economics, State University of New York, [SUNY]: Plattsburgh, USA

[2]Emeritus: Department of Accounting, School of Business and Economics, State University of New York, [SUNY]luskej@plattsburgh.edu : Plattsburgh, USA & Emeritus: The Wharton School, [Dept. Statistics], The University of Pennsylvania, USA. lusk@wharton.upenn.edu & Emeritus: Chair: International School of Management: Otto-von-Guericke, Magdeburg, Germany [https://www.uni-magdeburg.de/en/]

Correspondence: Mohamed Gaber, Department of Accounting, School of Business and Economics, State University of New York, [SUNY]: Plattsburgh, USA. gabermk@plattsburgh.edu

Received: July 28, 2023 Accepted: August 20, 2023 Online Published: August 24, 2023

doi:10.5430/afr.v12n3p69 URL: https://doi.org/10.5430/afr.v12n3p69

Abstract

Context The selection of the data-measurement metric should be the initial decision in the current field of Data Analytics [DA]. We discussed with DA-students and -colleagues re: vetting of the *Nature of the Data* to be used in decision-making. Uniformly, they indicated that ***rarely*** are vetting-screening-tests used for the accrued-data to determine if the *Nature of the Data* is in sync with the expected veracity of the DA-inferential results. This reticence seems to create inferential-issues that may well compromise the acuity and relevance of the inferential-output of DAs. To address these inferential-issues, we have selected a typical DA-screening *problématique.* First, we will vet the *Nature of the Data* and then address the *problématique.* Additionally, we will assume that the *problématique* requires the cooperative tri-interaction of: The Chief Operating Officer, The Financial Analysis Group & The Data Analytics Group. To illustrate these interactions, we suggest that Data Analytics Group is configured as "Internal Consultants" thus avoiding outsourcing elections. In this proposed Data Analytics Group-context, the Data Analytics Group elects to use a Factor Model [FM] as the Screening platform to "deconstruct" the Pearson Product Moment association profiles for the Data-Panels pursuant to addressing the *problématique*. ***Features***: We will detail: (i) A Data-Panel Screening protocol, (ii) A Factor pedagogic illustration, (iii) The Carvalho-script re: electing the Geometric-data-context, and (iv) The demonstration of these cooperative tri-cooperative interactions using the *Microsoft*™, *Inc.* [MSFT] Data-Panel. The overall goal is to offer illustrations, the intention of which, is to assist the pedagogical needs for instructors, and to populate the panoply for researchers and practitioners with effective and efficient inferential tools. ***Results*** We found that the Geometric-context was likely for our sample of market-traded organizations. Thus, we used the *ln*-transformation for the Panel of Data for our sampled firms. Additionally, we used a Factor Model as the screening tool for addressing the selected the *problématique*. The Data Analytics Group was cast as an "InSource" as this seems to be current institutional configuration adopted by many MNCs.

Keywords: VariMax Rotation, Eigenvalues, [EBIT, EBITA & EBIT], Principal Components

1. Introduction

1.1 Overview

In the past few years most academic institutions of higher learning have created programs, the nature of which, is to focus on using analytic tools to create valuable intel by processing data to aid managers in their current decision-making *milieu* re: the allocation of institutional resources. For example, the State University of New York: Plattsburgh has created a Data Analytics Master's degree in the Management, Information Systems & Analytics [MISA] Department. The Undergraduate and the Master's Degrees are among the most popular student elections in the MISA-Department as they open access to corporate professional engagements in Data Analytics Departments & Divisions. Thus, the emergence of Data Analytics in academia as driven by the need for DA-expertise in the corporate world reinforces the very standard and the very critical ***prime-directive*** of the *Analysis of Data* that pre-dates ***Data Analytics*** by decades—to wit:

> ***The prime-directive re: Data Analysis is to make sure that the statistical data-transformation protocol, if any, and the data-analytic-protocol under consideration are in sync with the goals of the analysis.***

Focusing on a related Micro-issue, we offer the following reflections from the introductory comment offered by Carvalho (2016), that we have found is not infrequently the case in the USA and in the Euro-Zone:

"How well have we been teaching arithmetic, harmonic, and geo-metric means to our students? In a recent article by C. R. Rao and colleagues (Rao, Shi & Wu 2014), we read

> '*Although the harmonic mean (HM) is mentioned in textbooks along with the arithmetic mean (AM) and the geometric mean (GM) as three possible ways of summarizing the information in a set of observations, its appropriateness in some statistical applications is not mentioned in textbooks.*'

1.2 The Prime Directive:

The Details One can think of Data Analytics as a Protocol [**DAP**] that is created to provide well-reasoned-intel that can be useful in informing the decision-making process. Thus, a DAP is unique to the particular *problématique* under investigation. A *problématique* is a System's concept often used by Russ Ackoff Note[1], here re-defined in a Data Analytical context, as:

> *An issue identified by members of a firm, the nature of which, requires the creative use of various mathematical, statistical and/or judgmental protocols needed to inform the decision-making process so as to control the issue in a limited temporal time-frame.*

In this sense, the DAP must be parameterized ***to suit the needs of the investigation***. This DAP-parameterization begins with examining the *Nature of the Data* that will be used in arriving at the needed decision-making intel. This is to say that the Prime Directive for any and all DAPs is to first determine the *Nature of the Data*; further, the usual analytical-driver, among many that are possible given any data-set, is intel with respect ***to the central tendency of the population from which the sample was randomly drawn.*** Thus, the *Nature of the Data* may be used as the parameterization guidelines of the DAP.

1.3 The Central Tendency of Choice

We have **s**elected the Mean of the population (μ) as the parameter of interest for our study. This is more or less the statistical-*Passepartout* in most analytical-contexts where sampling is the data collection modality. Consistent with the suggested disconnect noted by: Rao, Shi & Wu (2014), we will examine the three usual Means in the statistical-*milieu*: The Arithmetic Mean [AM], The Geometric Mean [GM] & the Harmonic Mean [HM], each of which assumes a very different generating-function for the data. These taxonomic choice-sets are presented by Carvalho (2016) as follows:

Table 1. Carvalho Mean Groups *These platforms are in most versions of EXCEL[*Microsoft™.Inc*].

Nature of the Data	Preferred Data Measure
Any Values in the Real Domain, where there is NO association among *any* of the point-values in the randomly selected data-set.	***Arithmetic Mean*[*AM*] Excel[AVERAGE]***
Any Values in the Real Domain > 0, where there IS association among *any* of the point-values in the randomly selected data-set.	***Geometric Mean*[*GM*] Excel[GEOMEAN]***
Any Values in the Real Domain > 0, where these values represent Rates or Ratios.	***Harmonic Mean*[*HM*] Excel]HARMEAN]***

Point of Information We are considering only the AM and the GM; the Harmonic Mean uses transformed or benchmarked-ratio-data and it is very often the case that the size of the Panel of these ratios is very small and is in the nature of Index-valuation. Also, mathematically, *when the GM and the AM can be calculated*, the HM can be formed as: HM= $\frac{GM^2}{AM}$. ***Link of Interest*** We recommend the following Link for more information on the AM, GM & HM. Note [2]

1.4 The Focus of this Research Report Given the introductory focus of this research report re: *The Nature of the Data* as the Prime Directive in a DA-context, following we will:

I. Assume a corporate decision-making *problématique*, the nature of which, requires coordinated-input from: The COO, The Financial Analysis Group in the Finance Department, and individuals in the Data Analytics Division [DA-Division],

II. Consistent with our consulting experience and the current trends in the advisor-sector, the DA-Division is now often configured as an intra-organizational consulting Division—sometimes referred to as an "InSource", i.e., a line-support group, usually reporting to the CFO. We adopt this advisory organizational configuration as an important aspect of this research report. Note [3] We offer suggestions as to how this DA-Division should interact in satisfying their in-house advisory *reason-d'être*,

III. We have selected a *problématique* that is ideal for demonstrating the utility of a Factor Analytic DA-approach. Our intention, in this regard, is to draw attention to the Factor Analytical model as a simple, effective, and under-utilized data-screening technique that is, in our experience, rarely used in problem solving situations. Thus, we hope to re-energize interest in the Factor-Model through its use in dealing with our selected *problématique*. We will introduce the Factor-model using the excellent Green, Tull && Albaum (2009) instructional discussion, finally

IV. We will illustrate most all of the intel needed to address the *problématique* by detailing the related computations. We hope that this will aid individuals in instructing these topics.

2. Instructive Creation of a Data Analytics Protocol [DAP]: The Interactive Process: Alpha to Omega

2.1 Advisory Services Imperative: The Next Evolutionary Phase Czerniawska (2005) details the evolution of various consultancy or advisory service organizational configurations; she notes that the time-line offers three configurations: Traditional, Integrated & Integrators. She notes (p.4) that integrators:

> "*have a wide range of skills internally but also subcontract to, or partner with, other suppliers in order to deliver an end-to-end service.*"

In the Integrators-class are Transaction & Transformational firms. she notes (p. 6):

> "*Transaction firms are therefore much more geared toward implementation and delivery than are relationship or even product firms.*
>
> *Transactions have four drawbacks from the consulting firm's point of view:*
>
> ** They are almost exclusively either consulting, technology, or outsourcing deals.*
>
> ** They are one-off deals that effectively focus on moving from A to B and then stopping; steady-state rather than continuous improvement.*
>
> ** Payment is based on the completion of the transaction (timely delivery, for example), rather than on business outcomes.*
>
> ** They reduce the opportunities to establish a long-term relationship with a client, which allows the consulting firm to stay in touch even while not carrying out work for the client*."

Most of these features in the Transactional Model lead to the next Phase: The Transformation Model. Czerniawska notes (p. 6):

> "*Two firms—IBM and Accensors—are unquestionably leading the way in changing this model, shifting their attention transaction to transformation - - - It is an expanding market, driven by clients who want to achieve a radical improvement in performance and who realize that they need a combination consulting, outsourcing and technology services to help them do this.*"

The Transformational Model precipitated the current trend to create Internal Advisory Services Divisions [IASD] that are Line-Divisions that report to the Chief Financial Officer or Chief Operations Officer. The IASD that offers advisory-support on organizational efforts to maintain effective, efficient and sustainable systems re: The conversion of institutional resources in support of the mission of the firm. An important critical benefit from a well-organized IASD is the *Flow-Event* proposed by Csikszentmihalyi (2000) that is an aesthetic noted as:

2.2 The Flow-Context The *Flow-Event* embodies and creates a "*holistic sensation that people feel when they act with total involvement*." (p. 36)

Simply, it seems that the IASD brings the members of the Firm together in a way that the Transaction or Transformation Models may not be able to do as the Advisory-activities are proposed by individuals ***external*** to the firm and so may work at cross-purposes to the Organizational Culture re: The problématique under investigation.

As a temporal-reality check, the IASD has existed since the 1960s when many organizations first installed Main-Frame computational devices. These firms very quickly realized that there were constant: non-trivial: maintenance, programming and trouble-shooting activities that were needed. *Voilà* the creation of the IT-Division of the Firm.

3. The DAP-advisory Context: A Suggested Configuration

3.1 The DAP-Launch The initiation of the creation of a DAP is a ***Request*** forwarded to the Section Leader of Data Analytics Division, Anaheim, CA by the Financial Analysis Group, Stamford, CT pursuant to the charge crafted by the COO Headquarters Division, Frankfort, Germany to create a Decision Analytic platform to:

> *Form a Portfolio of: Trading-market Investment Opportunities, the nature of which, are Flexible Short-term Cash Swap possibilities that are likely to outperform the interest-bearing instruments that we have been using over the last few years.*

The next set of DAP-interactions are taken by the Data Analytics Division and are presented following.

Data Analytics Division The following are the likely steps taken to create the **DAP[Fin[CashSwap]]**.

After a Zoom™ meeting with the Finance Division, a set of 21 Firms collected by the Finance Department that have been market-leaders for many years, were forwarded to the Data Analytics Division. Note [4] Each firm, had Panels of 13 Years of data. The variables of interest selected by the Finance Group were:

Four Profile Action-Variables: {EBITDA, *EBITDA_MARGIN[%]*, EBITA & EBIT}, and

Three *Results Measures*: {*GROSS_MARGIN[%]*, *OPER_MARGIN[%]* & *PROF_MARGIN[%]*}.

These seven-variables were obtained from *Bloomberg*™:[Income Statement[GAAP] for each of the 21-Firms. Their Bloomberg[BICS:Tickers] are noted in Appendix A.

3.2 Initial Parameterization of the DAP[Fin[CashSwap]] The Prime Directive is to select the data-measure most amenable to creating valid inferential information. This requires the determination if there is evidence of Panel-point-association over the 13-annul values for each of the 21-firms. A germane testing-measure in this regard is the: Longitudinal Pearson Product Moment [PPM]-association of the 13-Panel points with the generalized Time-Index. Specifically: In *EXCEL[Script]* for a Firm for Variable [k]:

$$PPM_{V[k]} = \text{CORR[Panel[n=13], TimeIndex[i, \{1, - - -, 13\}]}$$

In this case, the Data Analyst produced the following profile:

Table 2. PPM-Profile for the 21-Firms Forwarded by the Finance Division

CORRelation Profile	Average & Range	Median
Firms, n=21	[55.7%] [Min=15.2%: Max =92.3%]	59.6%

Discussion : The Computational Codex The DA-Group computed the CORR for each Variable for each Firm. Then the AVERAGE of these seven-Variable-PPM for each Firm was computed. This gives 21-Firm profiles that are reported in Table 2. ***Results*** There is clear and almost uniform evidence that there is pervasive PPM-association among the Panel-Points for the Firms. ***Action Implication***: The inferential measure ***required*** in this context is the Geometric Mean as the central-tendency inferential measure.

3.3 Transforming the BBT-Downloaded Firm-Data This ***Action Implication*** requires that the Geometric-context govern the inferential-montage of the DAP[Fin[CashSwap]]. This being the case, and given the Carvalho (2016, p.270) script, the transformation of the BBT-data needed to create the Geometric-context is the log-transformation of each datapoint in the Panel of the BBT-Firm datasets. Carvalho indicates that the Geometric- Mean [GM] is:

$$GM \equiv EXP[[1/n] \times [Sum[ln[x_i] \text{ i:1, - - -, n}]] \qquad EQ1$$

where: *ln* is the Natural log [*ln*] of the BBT: n-Data-points[n=13]: [*x*]s, EXP is the Exponential inverse of the Natural log[LN], and ≡ indicates that that the Right Hand Side of EQ1 is the computational definition of the GM. For example, using the BBT dataset for *Microsoft*™Inc. Appendix[**B2**]MSFT[EBITDA_MARGIN[%]], the Geometric Mean is:

$$GM[38.54] \equiv EXP[[1/13] \times [47.473]]$$

$$GM[38.54] \equiv EXP[3.65]$$

As a transformation software-check, the EXCEL[GEOMEAN] is:

$$[PRODUCT[Appendix[\textbf{Panel[B1}]]]^{\wedge}[1/[COUNT[\textit{Values in Panel[B1]]}]$$

$$GM = [[4.1436E20]^{\wedge}(1/13)] = 38.54$$

Point of Information: Cascading Implications Following the Prime Directive, we are obligated to use the Geometric-context for the datasets to address the charge of the COO. See Tables 1 & 2. Further, this has implications for the nature of the data that can be processed in the Geometric-context; *to wit*, all the Panel-Points must be > 0. *Continuing*: The implication of this is that of the 21-Panels that were sent by the Finance Group, only 13-firms had qualifying Panel-Points. See Appendix A. A *Zoom* meeting between the Finance Group & The Data Analytics Division confirmed the acceptability of paring-down of the test-firms selected by the Finance Group. Thus, as the *Nature of the Data* has been ascertained and the qualifying firms identified [Appendix A], the next phase is: The selection of the ***Data Processing Model*** for the execution of the DAP[Fin[CashSwap]].

4. Model Selection: Probing the Cash Swap Milieu

4.1 Overview

There are a vast number of models, the nature of which, could be parameterized to create valid- and useful-intel required to inform the decision-making requirements as expressed by the COO's charge. In this regard, there was a *Zoom*-meeting among three of the senior Modeling Experts of the Data Analytics Division & the Finance Group. After the *Zoom*-discussion, as closure, Tameka Walker, PhD, the DA-section-leader, recommended a *Factor Model* to create the required information needed to inform the Cash-Swap decision-making process. In this regard, she suggests a pedagogic-illustrative Factor Model to ascertain if the Finance Group would accept a Factor Model [FM] as their Screening Model in executing the DAP[Fin[CashSwap]]. Note [5]

4.2 Zoom-Meeting Subject: Details of the Factor Model

The Open-Access-[Fin]-link *Zoom*-meeting was booked for three hours. Preparatory to the *Zoom* meeting. the following basic information to be discussed by the DA:Zoom-Leader, was uploaded to the Finance Group-ShareSpace.

4.2.1 Instructive Illustration of the Factor Model

The following illustrative Factor Model [FM] example was liberally paraphrased from Green, Tull & Albaum(2009) to elucidate how the FM works. Note this is the GTA-Example:

Assume that ***Space Aliens*** land on Earth circa 2023. They send colleagues to investigate the nature of various aspects of life on Earth. One group finds, on Amazon™, technical paraphernalia called ***Computers***. To investigate or deconstruct or understand the Nature of these Computers, they randomly purchase **15**-Computers and for each they carefully examine: (i) The Amazon-site where the manufacturing specifications are detailed, (ii) Technical Articles and Advertisements on Computers, (iii) History Books on Computing Technology & (iv) Consumer Ratings from Technical Journals. This yielded **7**-descriptors or variables that the Aliens intend to use to understand the Nature of Computers. Specifically, they selected the following variables*: (1) Basic Processing Speed [Addition], (2) Advanced Processing Speed [Multiplication], (3) Minimum Storage, (4) Maximum Storage, (5) Add-Ons Non-buffer Capacity, (6) Cycle Time & (7) Diagonal Screen Size: [30cm : 46cm]*. This produces a Factor Input Matrix of size [**15** × **7**] that is the data-input to the FM. **Note** In this case, we are not permitted to display this Matrix as it is the copyright policy of Green, Tull & Albaum [Prentice-Hall]. However, we added Variable (7) to the six used by GTA. This creates slightly different values in the Factor-Profile. Thus, we are permitted to produce and display the Factor Analysis Profiles.

At this point, we will describe the Factor Model in very ***General non-technical terms***; this is needed as the Factor Model is driven by very complicated Mathematical and Statistical operating systems.

4.2.2 The Factor Model: A ***Simplified*** Descriptive Version

Note, we are using the GTA-Computer example. Assume that the Aliens have invented a Linguistic Concept Synthesizer [LCS] detailed as:

> ***Phase A*** *The* Linguistic Concept Synthesizer [LCS] *is a software-driven model, into which, an analyst inputs the* Factor Input Matrix of size [**15** × **7**] that are the **7-*descriptions*** *of the operational aspects of the* ***15-technical apparatuses****. Where, the LCS displays* ***7****-indications of the explanatory power of* ***EACH*** *of the 7-possible Factors presented individually. With this Factor-information, the analyst evaluates this Explanatory Factor-Profile and selects the number of Factors [1, 2, - - -, or 7] that seems to provide the best information for deconstructing the Nature of Computers.*
>
> ***Phase B*** *After the analyst makes the decision as to the number of Factors to profile, then the LCS produces a matrix, the size of which is:* size [**7** × **k**], where: **k** is the number of Factors selected by the analyst. This is called the Factor-Profile.
>
> ***Phase C*** *Then the LCS populates the Factor Profile by producing* ***intra****-factor evaluation weights for the* ***7****-variables for the* ***k****-Factors. These weights are used to Deconstruct/****Name*** *the Nature of the functionality of these* ***15****-apparatuses. Thus, the purpose of the LCS is to allow the analyst to select from among these* ***7****-iterated descriptive Profiles, a particular Profile that the analyst judges to be the most informative so as to better understand the Nature of the overall functionality of these technical apparatuses.*

A graphic of the LCS is presented in Figure A.

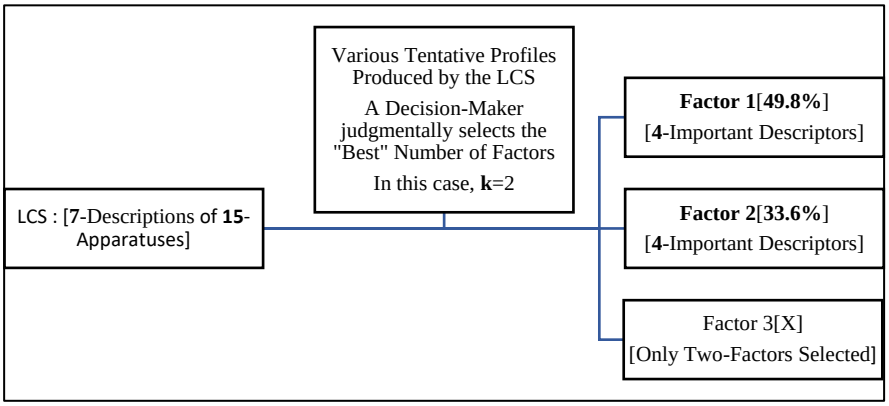

Figure A Overview of The Three Phases of the LCS Factor Model: using the GTA-pedagogic-illustration

Discussion Assume that the GTA-Factor Input Matrix of size [**15** × **7**] was inputted to the LCS software, and during the Phase B the GTA-analyst selects TWO Factors. The explanatory weights are noted in Figure A as: Factor 1[**49.8%**] & Factor 2[**33.6%**] where the explanatory power of the TWO-Factor solution is **83.4%** [49.8% + 33.6%]. Finally, the LCS produces the GTA-augmented Factor Profile following:

Table 3. Factor Score Weighs for the GTA-Computer Illustrative Example

7-*descriptions of GTA- Illustration*	**Factor 1 [49.8%]**	**Factor 2 [33.6%]**
Processing Speed [Addition]	**0.9464**	-0.2477
Processing Speed [Multiplication]	**0.9580**	-0.1863
Minimum Storage	-0.0960	**0.8849**
Maximum Storage	-0.0690	**0.9376**
Add-Ons Non-buffer Capacity	-0.1840	**0.7223**
Cycle Time	**0.9711**	-0.0746
Diagonal Screen Size: [30cm : 46cm]	**0.5050**	**0.4379**

4.2.3 Discussion Table 3

The GTA-analyst, based upon the ***7-indications of the explanatory power of EACH of the 7-Factors***, selected the Two-Factors Model that has the *explanatory power of*: **83.4%** [49.8% + 33.6%]. Then, the LCS-software produced information presented in Table 3. In this case, the Factor Score Weight-values are akin to PPM-Correlations and are produced by the LCS-software. We will elaborate on these Factor Score Weights in the next section of this research report. We have bolded those Descriptor-scores that are "meaningful" in Naming the Two-Factors—this is the deconstruction or **Naming** Phase of the LSC-Model.

For Factor 1, there were **4**-variable descriptors that were important in Naming Factor 1,

For Factor 2, there were **4**-variable descriptors that were important in Naming Factor 2, and,

For Factor 3, was not selected as an important aspect of the GTA-Example.

Thus, ***judgmentally***, the analyst gives a linguistic interpretation to each Factor—***to wit, the analyst Names the Factors***. For example, for the GTA-example,

For Factor 1 there were four (4) Descriptive-Variables that the LCS model grouped together as being "Meaningful Associated as **bolded**:

Processing Speed [Addition]

Processing Speed [Multiplication]

Cycle Time, and

Diagonal Screen Size: [30cm : 46cm]

The explanative power for Factor 1Named:[***Speed***] was **49.8%**

For Factor 2 there were four (4) Descriptive-Variables that the LCS model grouped together as being "Meaningful Associated as **bolded**:

Minimum Storage

Maximum Storage

Add-Ons Non-buffer Capacity, and

Diagonal Screen Size: [30cm : 46cm]

The explanative power for Factor 2 Named:[**Storage *Capacity***] was **33.6%]**

Discussion The LCS Scoring offers, as a vetting-reality, ***what we already know about Computers***. They are computing-apparatuses that have Two dimensions or Factors: ***Factor 1 [Speed]*** & ***Factor 2 [Storage Capacity]***. As an additional variable, both are associated with Screen Size. In the PPM-correlation mode: The More the Storage Capacity & The Higher the Computational Speed, the Larger is the Display Screen. ***Summary[GTA]*** This is a clear indication that the LCS[Factor Model] reports ***The Obvious***; this is ***A Vetting-Replication of a Generally Accepted Reality***—it is a generally accepted inferential vetting technique.

4.2.4 Summary: Factor Model

This is the end of the information that was uploaded to the Financial Analysis Group re: ***The Screening Model Selection***. After the Zoom-meeting of the DA-Presentation of the FM, there were four 90-minute Q&A sessions offered by the DA-Group over a one-week period. Finally, after the Q&A-discussions among the Finance-Group & the DA-Group, the Financial Analysis Group decided to accept a Factor Model for the project: DAP[Fin[CashSwap]]. In this case, the next phase proposed by Dr. Walker is an illustrative DA-technical analysis of one of the firms selected by the Financial Analysis Group. She chooses: *Microsoft*™, Inc. [MSFT]. Following is a DA-Group ***proto-presentation*** of MSFT; a *proto-presentation* focuses on the technical nature of the Factor-deconstruction and is only an illustrative analysis of the logistical -aspects of the deconstruction of the FM for MSFT. In this regard, as the DA-Group, we will offer three DA-Deliverables, the intention of which are to provide examples of the DA-investigation that may aid the Financial Analysts in formulating their investigations re: Cash-Swap targets. Following are**:**

I. Deliverable **A** The ***Deconstruction*** of the MSFT: Factor-Results of the FM,

II. Deliverable **B** The Classification of the ***Investigated Potential*** of the Firms Accrued by the Financial Analysis Group, and

III. Deliverable **C** The ***Illustrative Computations*** and Intel of the Geometric-Context.

5. The DAP[Fin[CashSwap]]: An Illustrative Analysis Enriched with the Computational Details

*5.1 DAP[Fin[Cash-Swap[**Deliverable A**] Using the FM-Factor-Screen for MSFT.* Recall, the seven variables taken from Bloomberg[Income Statement[GAAP:**IS**]] Note [6] are:

Four IS-Profile Action-Variables:{*EBITDA, EBITDA_MARGIN[%], EBITA & EBIT*} and

Three IS-*Results Measures*: {*GROSS_MARGIN[%], OPER_MARGIN[%]* & *PROF_MARGIN[%]*}.

*The AIS-Definitions of these variables downloaded from Bloomberg[15June2023] are presented verbatim in Appendix C. **Note:** The Bloomberg reporting protocol for Percentages is: ##.##. There is NO affixed %; so this appears as a number in Bloomberg downloads.

As discussed above, after testing the longitudinal-PPM Correlation, we assumed the data is produced by a Geometric Generating Process. As is the case for all Factor Models, there is information provided as to the information-content for each of the possible factors. In the FM, this information is provided by the Eigenvalue Profiler. The function of this Eigenvalue Profiler is to provide intel critical in selecting the ***Number of Factors*** to be displayed for the Factor ***Results*** Profile. **As implied in the LCS[GTA]-overview**, selecting the number of Factors is the intel that informs the decision-making process; however, the technical aspects underlaying the Eigenvalue Profiler are mathematically and statistically beyond the scope of this research report. However, a simplified-version of the functioning of the Eigenvalue Profiler, as an elaboration to the LCS-discussion, is critical to understanding the FM.

5.1.1 The Eigenvalue Profiler

The FM uses the Pearson Product Moment [PPM]-Correlation Matrix or the Co-Variance Matrix of the variables as its primary input. Both of these inputs give essentially the same results in most instances; for consistency, we will use the Eigenvalues of the PPM Correlation Matrix. The FM assumes initially that the number of Descriptive Variables define the dimensions of the Variance-Axis-Space. For MSFT, the Variable Space has Seven-Axes, where initially each Axis accounts for its-own-Standardized OLS-Variation. Thus, this gives the genesis of the Term: ***Eigenvalue*** which comes from the German where it means: ***Its Own Value***. For MSFT, for the Seven-variables entered, there are, in total, Seven-standardized units of variation or 1.0 Variance Unit *per* each of the Seven Descriptive Variables; this accounts for 100% [7/7 × [100%]] of the total variation of the PPM-Correlation Matrix. The critical decision-intel of the Eigenvalue Profiler is: ***A particular Eigenvalue gives the proportion of variation explained for the re-oriented projection of the variance based upon the PPM-correlation among the seven-variables.*** Thus, the Eigenvalue Profiler reports the amount of variation explained after the rotation—called a Veri-Max-Factor Rotation of the Principal Components. For example, using the LCS[GTA]-case, the Basic Processing Speed Eigenvalue Loading-Weight was: 3.4838. This can be transformed to the projected explanative power of Factor 1 as: 49.8% [3.4838 / 7].

5.1.2. The Factor Profile & Naming the Factor

After the analyst selects the number of factors to be displayed in the Factor Profile, the FM produces the Factor Profile where the importance of each of the Descriptive Variables is reported for each of the Factors. This is referred to as a Factor Variable ***Loading-Value***. For example, for the LCS [GTA]-Example, for Factor 1 for the Descriptive Variable: ***Basic Processing Speed***, the Factor Variable ***Loading-Value*** weight assigned by the FM is **0.9464. The importance of each of these Descriptive Variables** depends on the judgmental of the analyst. A computation that is often used to identify "Important Descriptive Variables" is to compute a heuristic False Positive Error [FPE]-Null for the Factor Variable ***Loading-Value*** weight. If this FPE[Null] is < 0.10, this usually rationalizes identifying that Variable as an important Variable-Descriptor for Naming the Factor. This computational protocol is presented following:

Computational Indication: The ***Factor Loading*** Value for Basic Process Speed is [**0.9464**], the heuristic p-value for the directional FPE[Null] of 0 is:

$$t = 10.5 = [ABS(0.946) \times (15-2)^{\wedge}.5] / (1-(0.946)^{\wedge}2)^{\wedge}.5$$

T.DIST.RT(10.5,15, (15-2)) is: [FPE] p-value of <0.0001; this is < 0.10

Implication If a particular Factor Loading on a Descriptive Variable has a FPE[p-value] that is < 0.10, then this is a likely indication that this Descriptive Variable is an important aspect of that Factor. The FM does not make this inferential calculation. We suggest creating a VBA-program to produce these inferential screens as they provide useful intel in collecting the information needed to Name the Factors.

5.2 The MSFT ProtoAnalysis The DA-Group's recommendation is to: *Select the Factors for rotation corresponding to the number of Eigenvalues for which the cumulative explanation is*: *Min*{(i) *At least 75% & (ii) There are at least Two Factors*}. The *Min*-Condition Factor Selection Rule [Min-CFSR] indicates the ***Smallest*** Number of Factors for which the *Min*[Condition] is TRUE. For example, for the MSFT, the eigenvalues are presented Column 1 in the following table:

Table 4. Eigenvalue Profiler[MSFT] DA-proto-illustration *We have rounded the Eigenvalue to four decimal-places.

Index Number & [Eigenvalue*]	Percentage Variance Explained	Cumulative Variance Accounted	Before Rotation Cumulative Percentage	
1 [3.7168]	53.097	0.5310	0.142857	[1/7] × 1
2 [2.9131]	**41.615**	**0.9471**	**0.285714**	**[1/7] × 2**
3 [0.2675]	3.822	0.9853	0.428571	[1/7] × 3
4 [0.0924]	1.321	0.9985	0.571429	[1/7] × 4
5 [0.0098]	0.14	0.9999	0.714286	[1/7] × 5
6 [0.0004]	0.005	1	0.857143	[1/7] × 6
0 [0.0000]	0	1	1.000000	[1/7] × 7

Discussion For the MSFT-dataset, after the Vari-Max-orthogonal-Principal Components-rotation of the MSFT's PPM-Correlation Matrix, there is a considerable consolidation of Variance of the Seven-Variables on the first two-Factors. Specifically, for the first Factor Axis **53.097%** [3.7168/7] of the common variance after rotation is reported for the first Factor; then, adding a second Factor-Axis **41.615%** [2.9131/ 7] of the common-variance explained is added. Thus, for the first two Factors {A & B} there is, in total, **94.712%** [[(3.7168) + (2.9131)]/7] of the common variation that is explained for the seven-variables. In this case, given our Min-CFSR] a TWO-FACTOR-structure was selected. In addition, sometimes it is helpful to compute a baseline measure called the ***Non-rotated axis Measure***. This is presented in Column 4. This is the explained variance if every variable was zero-PMM-co-correlated. For example, the two-Factor score, if all of the Seven-variables were un-PPM-correlated. In this case, the rotated-Principal Components Two-Factor explains ***94.712%*** of the variance while the Unrotated-Space accounts for ***28.571%***—this is, of course, a convincing indication that the two-Factors as rotated are very likely to be informative.

5.2.1 Evaluating the Loading Values for the MSFT re: The Cash Flow Swap

The next phase presented by the DA-Group is to examine this Two-Factor Loading-Profile for MSFT as outputted by *SAS*™[*JMP*™[*Multivariate*]v.13]]: *Factor Analysis*. See Table 5 following:

Table 5. MSFT Panel Factor using ln-transformed BBT-reported-values

Firm [MSFT]BBTs	Factor A	Factor B	Rotation Proof
EBITDA[IS-Value]	**0.9982**	-0.0201	1>0.99
EBITDA_MARGIN[%]	**0.4728**	**0.8481**	1>0.94
EBITA[IS-Value]	**0.9915**	0.1110	1>0.99
EBIT[IS-Value]	**0.9691**	0.2147	1>0.98
GROSS_MARGIN[%]	**-0.6848**	**0.6806**	1>0.93
OPER_MARGIN[%]	0.0361	**0.9851**	1>0.97
PROF_MARGIN[%]	0.0020	**0.8977**	1>0.80

Details: Table 5 The Factor Loading Scores are usually given a descriptive context. The DA-Group offers to the Finance Group the following linguistic-codex for better understanding the intel of the MSFT-Factor Table:

5.2.2 Linguistic Taxonomy: Suggested by the DA-Group:

Table 6. Linguistic Code proposed by the DA-Group *Harman Unique Loading Value is: ABS[$\sqrt{0.5}$] = 0.7071

Factor Loading Screen	Linguistic Indication	Codex
[> 0.7071* to 1.0]	**Unique Positive Factor Defining Variable**	[U+]
[> 0.50 , . . . , ≤ 0.7071]	**Dominate Positive Factor Defining Variable**	[D+]
[≥ 0.25 , . . . , ≤ 0.50]	**Suggestive Positive Factor Defining Variable**	[S+]
[> (-0.25) , . 0 . , <0.25]	**A Non-Factor Defining Variable**	[NonF]
[≥ (-0.50) , . . . , ≤ (-0.25)]	**Suggestive Inverse Factor Defining Variable**]	[S–]
[≥ (-0.7071) , . . . , < (-0.50)]	**Dominate Inverse Factor Defining Variable**	[D–]
[-1 to < (-0.7071)]	**Unique Inverse Factor Defining Variable**	[U–]

In the case for MSFT, the Linguistic Codex and Highlighting scripted is:

Table 7. Final edited Factor Profile for MSFT

Firm [MSFT]BBTs	Factor A	Factor B	Rotation Proof
EBITDA[IS-Value]	**[U+]**:0.9982	[NonF]-0.0200	1>0.99
EBITDA_MARGIN[%]	**[S+]**:0.4728	**[U+]**:0.8481	1>0.94
EBITA[IS-Value]	**[U+]**:0.9915	[NonF]0.1110	1>0.99
EBIT[IS-Value]	**[U+]**:0.9691	[NonF]0.2147	1>0.98
GROSS_MARGIN[%]	**[D–]**:-0.6848	**[D+]**:0.6806	1>0.93
OPER_MARGIN[%]	[NonF]0.0361	**[U+]**:0.9851	1>0.97
PROF_MARGIN[%]	[NonF]0.0020	**[U+]**:0.8978	1>0.80

5.2.3 MSFT: Discussion

In this case, we have coded the **NonF** as shaded and the Loadings that have PPMFPE[p-values] <0.10 as **Bolded**. Others may use different highlighting schema to draw attention to the salient aspects of the Factor-Profile. Often, color-coding seems to be the highlighting-mode. The inferential profile as it pertains to the DAP[Fin[MSFT{CashSwap] is presented as a detailed proto-deconstructed-analysis of MSFT. ***Point of Protocol*** All the following information for ***Deliverable A***, as it pertains to MSFT, is *ONLY a DA-Group proto-example of an analysis **that may be** conducted by the **Financial Analysis Group***—the group charged by the CFO & COO with making the decision re: Selecting Viable Cash-Swap targets. As is evident, there are numerous relationships that ***could be*** profiled in the FM[MSFT]-Two-Factor Profile of Table 7. We will treat two analyses in suggestive detail and suggest two others at the conclusion of the Deliverable A sections for both Factors A & B.

5.3 Judgmental Profile[MSFT[Deconstruction Judgmental Profile]]:

Deliverable A For **MSFT[Factor A]**, given the significance of the Variable-Loadings, the name assigned by the DA-Analyst is: [**The IS-Earnings[Revenue] Block]**. *Rationale*: The Earnings variables: {EBITDA, EBITA & EBIT} are very highly co-Factor-associated; all have p-values <0.10. Basically, their Income Statement-association-values are effectively identical re: their change-profiles over the 13-Panel points. Additionally, the loading-association of this **IS-Earnings[Revenue] Block** with the EBITDA_MARGIN[%] is [S+] indicating that the ratio of: [The trailing 12-month EBITDA] divided by [The Trailing 12-month Sales] basically moves directionally with {EBIT, EBITA & EBITDA}. This is likely a desirable indication in that as **Earnings** increase over the annual tracking period so also does the EBITDA_MARGIN[%]**.** ***However,*** these four **IS-Earnings[Revenue] Block**-variables link-up in an interesting way with GROSS_ MARGIN[%] where there is a [D-] inverse association. This suggests that as Earnings Increases the GROSS_ MARGIN[%] measured as: [(Net Sales - Cost of Goods Sold[COGS]) * 100 / Net Sales] decreases. Perhaps, this indicates the ***usual returns to scale*** where higher earnings result in COGS increases at a higher rate than Earnings at the margin thus resulting in an inverse PPM-associational relationship between: Earnings & GROSS_ MARGIN[%]. These Factor A-relationships, as they exist in MSFT, suggest that in analyzing the quality of MSFT as a Cash Swap, the Finance Analysis Group should pay attention to the very dominant

and highly positively associated EBIT-Block that acts as a driver for the compromised or inverse effect GROSS_ MARGIN [%]. The reason to give pause relative to the Cash-Swap potential of MSFT is that at some point in the future there may be "tipping" point where the Earnings-Block may be compromised by the shrinking GROSS_ MARGIN [%] as the COGS move to point where it adversely affects the EBIT-Block. If traders in the market perceive that this tipping point is possible *or eminent* that could compromise the potential of using MSFT as Cash-Swap.

5.3.1 Alternative Investigative Analysis Accounts

There are two-interesting investigative-aspects that may be gleaned from Factor A:

I. Factor A as defined, has no loading association with: OPER_MARGIN[%] or with PROF_ MARGIN [%]. This indicates that Factor A is unrelated with the changes in OPER_ MARGIN[%] or with PROF_ MARGIN [%] over the Panel. This is indeed a curious relationship; actually, it requires a careful detailed analysis of MSFT to determine the projective implications for the stability of the Earning Block trajectory that would likely be affected if, over time, there is **NO** PPM-association of Earnings with OPER_MARGIN[%] or with PROF_ MARGIN[%]. These profile-dis-connects seem relatively anomalous and may be identified by individuals in the Market as such curious organizational features so as to create a certain AIS-perversity that could compromise MSFT as a Cash-Swap. This of course depends on the results of - Financial Analysts findings.

II. Another interesting feature of the significant and co-uniform-Loadings of EBIT, EBITA & EBITDA is that it is not likely that the MSFT-senior managers have used Discretionary Accruals re: Revenue or Deprecation or Creative Tax-Accounting and, the like, to differentially manage reported Earnings-Profiles over the 13-annual Panel-Points. This may be construed by traders in the Market as a positive indication of stability and thus engender confidence in the wisdom of MSFT-earnings-management actions. This may support -the use of MSFT as a Cash-Swap.

5.3.2 MSFT[Factor B] The DA-Analyst labeled Factor B: Percentage Co-Associational Results Block Simply, all ***four*** of the Margin-percentage-profiles: EBITDA_MARGIN[%] & {GROSS_ MARGIN [%]; OPER_ MARGIN[%] & PROF_ MARGIN[%]} are [U+ or D+]-associated over the Panel for Factor B. Using Factor A to benchmark Factor B, creates an interesting associational relationship re: the behavior of the co-association-loadings between Factor A and Factor B with respect to:

EBITDA_MARGIN[%] & GROSS_MARGIN[%]

Specifically, as an elaboration of the discussion of **Factor A IS-Earnings[Revenue] Block, f**or Factor A we have:

EBITDA_MARGIN[%] [**S+**]:0.4728

GROSS_ MARGIN [%] [**D–]**:-0.6848

while for Factor B, we find:

Panel B Percentage Co-Associational Results Block:

EBITDA_MARGIN[%] **[U+]**:0.8481

GROSS_ MARGIN [%] [**D+]**:0.6806

Note that all four p-values for the above Loadings were scored as: {[U+], [D±], or [S+]} and are <0.10. This profile is tantalizing and begs a deconstruction to discover if the relationships between Factor A & Factor B could impact the Cash-Swap desirability of MSFT. In this regard, consider the following:

The EBITDA_MARGIN[%], scored as [**S+**] & [**U+**] for Factors A & B respectively, is also paired with {Factor A[GROSS_ MARGIN[%] [**D–]}**:-0.6848 & {Factor B [GROSS_ MARGIN[%] [**D+]}**:0.6806} that have opposite indications relative to the definition of Factors A & B. This indicates a ***schizophrenic-profiling*** as follows: Factor A **Higher** EBITDA_MARGIN[%] drive **DOWN** the GROSS_ MARGIN[%] [**D–]**:-0.6848; while for Factor B, **Higher** EBITDA_MARGIN[%] drives **UP** GROSS_ MARGIN[%] [**D+]**:0.6806.

Implications Thus, EBITDA_MARGIN[%] is a "***bi"-driver*** and so can have two different Result-profiles. Specifically, the Finance Group needs to decide if in the short-term Cash-Swap time-frame, the Management of MSFT will create an action plan, the intent and success of which is to:

(i) increase Earnings in the ***three-EBIT-Group-variables***: Result→ GROSS_ MARGIN[%] expected to DECREASE, or

(ii) be addressed to increasing ***Returns at the Margin:*** Result→ GROSS_ MARGIN[%] expected to INCREASE.

Thus, in this case, the Finance Group must decide ***{IF*** & ***HOW*****)** individuals active in Market **{**will react to MSFT's action taken**}** and **{*IF*}** that projected action results in an **{Increase** OR **Decrease}** in the short-term trading-value of MSFT. Thus, there are four questions that the Finance Group may need to address so as to decide if MSFT will be a viable Cash-Swap target.

Also, to be clear, we at the DA-Group have selected, essentially, one-set of relationships that seemed to be of interest to us. There are more than a few relationships in the MSFT-Factor Results Table that may *tweak* & *pique* the interest of the Finance Group. For example, the relationship for:

Table 8 Alternative Curious MSFT Panel Relationships of possible investigative interest to the Finance Group could certainly be the subject of interest to the Financial Analysis Group.

Firm [MSFT]BBTs	Factor A	Factor B
OPER_MARGIN[%]	[NonF]0.0361	**[U+]**:0.9851
PROF_MARGIN[%]	[NonF]0.0020	**[U+]**:0.8978

*5.4 DAP[Fin[Cash-Swap[**Deliverable B**] Analyses of the 13-Panels: Inferential Scope*

In the MSFT case, the DAP-analyst decided to offer a **DAP**[Investigative Potentiality] analysis. For example, the Financial Analysis Group may be interested in the nature of the Factor-loading-profiles that may likely identify possibly interesting matches of the four EBIT-Variables with respect to the three Margin-Set[%]-Variables. Assume that the question offered by the DAP-analyst is:

For each of the 13-firms:

> *How many of the four **EBIT**-variables have directional p-values < 0.10 **and** How many of the three **Results**-Marginal[%]-variables have directional p-values < 0.10 for each of the HFM-Factors?*

5.4.1 Screening Rationale

This profile would enable the Finance Group to identify a rich set of firms that, upon further analysis, may prove to be excellent Cash-Swap short-term market targets. It is possible, but not usually the case, that for each of the two-factors identified by the Eigenvalue Profiler, there could be seven-variable post-rotation loading values, the p-values of which, could all be < 0.10. Thus, the max-profile for each of the two-factors {A & B} is 7 or 14 in total.

In this regard, for ALL 13 Firms, the Eigenvalue Profile using the PPM-Matrix indicated that two-Factors would be the suggested Factor Veri-MAX Rotation re: The Principal Components; thus, Table 9 presents for each Factor the number of Factor-Loadings, the p-values of which are, <0.10.

Table 9. Screened Firm relative to Investigative Potential

Firm [BBTs]	Factor A	Factor B	DAP: Suggestions
Coca Cola [KO]	6	4	Interesting[71.4%]
Pepsi [PEP]	7	7	Excellent [100%]
Microsoft [MSFT]*	**5**	**4**	**Interesting[64.3%]**
Apple [APPL]	4	5	Interesting[64.3%]
MacDonald's [MCD]	6	7	Excellent [92.9%]
Wall Mart [WMT]	4	3	Interesting[50.0%]
CVS Health Inc. [CVS]	4	3	Interesting[50.0%]
Bed Bath Beyond [BBBY]	4	4	Interesting[57.1%]
Hersey [HSY]	5	3	Interesting[57.1%]
Tiffany [TIF]	6	4	Interesting[71.4%]
Molson [TAP]	6	6	Excellent[85.7%]
Disney [DIS]	7	7	Excellent [100%]
Comcast[CMCSA]	6	3	Interesting[64.3%]
Average			**Interesting [71.4%]**

*Factor Results for MSFT are Bolded.

5.4.2 DAP: Codex for Table 9

The two Factors for each of the 13-Firms accrued by the Financial Analysis Group are evaluated for the number of the seven-variables that have p-values < 0.10. In Column Four, the DA-Analyst has given a linguistics-context to the richness of the investigative potential as follows:

Excellent: There are a large number, much more than a majority, of Variable-Loadings the p-values of which are < 0.10. ***Judgmental Codex***: If the Firm p-value-percent is: [≥ 80% : ≤100%] this FIRM is scored as offering an ***Excellent*** potential for a compressive and meaningful Financial investigative analysis. Summary Indication: ***Excellent*** For example, for Pepsi [PEP] there are seven-Loadings for Factor A and seven for Factor B that were < 0.10, yielding: 100% [(7 + 7)/14 × 100%], **OR**

Interesting: There are a number, a majority or more, of Variable-Loadings the p-values of which are < 0.10. ***Judgmental Codex***: If the Firm p-value-percent is: [≥ 50% : <80%] this FIRM is scored as offering a reasonable potential for a compressive and meaningful Financial investigative analysis. Summary Indication: ***Interesting***, **OR**

Poor: There are only a few, less than a majority, of Variable-Loadings the p-values of which are < 0.10. ***Judgmental Codex***: If the Firm p-value-percent is: [≥ 0% : <50%] this FIRM is scored as ***not*** offering a reasonable potential for a compressive and meaningful Financial investigative analysis. Summary Indication: ***Poor.*** **END**

5.4.3 Assessment of the Investigative Quality

The 13 Firms are scored using the consultative judgment of the DA-Group as to the quality of information that ***may be collected*** to arrive at a reasoned evaluation as to the desirability of the firm as a short-term Cash-Swap target. The DAP-analysis was based upon the number of p-values for the Variable-Loadings < 0.10. ***Summary:*** *The Profiles of Table 9 are most interesting.* ***All of the firms were suited to deconstructive analyses.*** As a sampled vetting of the intra-Variable Associations, we selected 28-Panels from our BBT-firms. We then created a matched random Sample-set of Panels using =RANDBETWEEN[Min, Max]. Then. We computed Cronbach's α for each of the 28-Panels blocked by the sampled firms. The Results were: BBT-Firms [Mean[0.74] & Median[0.79]] and for the Random-Panels [Mean[–0.16] & Median[–0.01]]. The acceptable/desirable ranges for Cronbach's α as reported by Tavakol & Dennick (2011) are in the range [0.70 : 0.95] also see: Nunnaly (1978). Using the Welch-ANOVA test the p-value

for the means that we found was <0.0001, indicating that the BBT-Sample and the Matched Random Sample are not likely to be samples from the same population.

*5.5 DAP[Fin[Cash-Swap[**Deliverable C**]:Computation Context of the Analyses of the 13-Panels*

Assume that due to the MSFT: ***bi-projection factor results***, the Finance Group selected the variable-pairs: MSFT[EBITDA_MARGIN[%]] & MSFT[GROSS-MARGIN[%]] for further analysis. As a preparatory step, the DA-Analyst computes the Geometric Mean and its 95%CI for these two-Panels to get range information for the Population Geometric Means of these variables for the MSFT-sample. This is a very standard profiling-step in creating investigative intel. These computations are not straightforward; to offer clarification, we will script-out these computations. Also, recall, to create the Geometric analytic context, the *ln*-transformation of the basic MSFT-Bloomberg download was taken. Thus, the MSFT, data as reported by Bloomberg, is re-stated as *lns* and so are all the Factor results. Thus, the Geometric Mean is the Arithmetic Average of the *ln*-transformed data. The 95%Confidence Intervals are easily found in Excel by using the, following platform:

Excel[Data[DataAnalysis[Descripivestaistics[95%ConfidenceIntervals]]]].

5.5.1 Blocked Investigative Analysis

Thus, for the *ln*-transformed data the Mean and the 95%Confidence Intervals are:

MSFT[EBITDA_MARGIN[%]]Appendix B[Panel 2]:

Mean [3.6518],

Lower 95%Limit [3.6518 – 0.0859] = 3.5659

Upper 95%Limit [3.6518 + 0.0859] = 3.7377

Where: The precision reported by the above Excel-platform is: *Confidence Level(95.0%)* = 0.0859.

MSFT[GROSS_MARGIN[%]]Appendix B[Panel 4]:

Mean [4.27699],

Lower 95%Limit [4.2770 – 0.0576] = 4.2194

Upper 95%Limit [4.2770 + 0.0576] = 4.3346

Finally, it is usually the case to re-transform the *ln*-parameters back to the basic MSFT Bloomberg data measures. In this case, one applies the ***EXP***-Excel function as follows:

MSFT[EBITDA_MARGIN[%]]:

EXP[Mean [3.6518]] = 38.5436 ≡ Geometric Mean,

EXP[Lower 95%Limit:3.5659] = 35.3702

EXP[Upper 95%Limit:3.7377] = 42.0017

MSFT[GROSS_MARGIN[%]]:

EXP[Mean [4.2770]] = 72.0240 ≡ Geometric Mean,

EXP[Lower 95%Limit: 4.2194] = 67.9927

EXP[Upper 95%Limit: 4.3346] = 76.2944

Discussion Given the Geometric Population-Context, there could be useful investigative intel to be found using Regression Projections.

5.5.2 Regression Projection

Assume that the Finance Group is interested in the forecasting expectation of MSFT's GROSS_MARGIN[%] ***if*** the EBITDA_MARGIN[%] were to be 5% above the High side of the 95%CI[EBITDA_MARGIN[%]]. The question of investigative interest is:

Is MSFT's projected GROSS_MARGIN[%] ***IN*** the 95%CI[GROSS_MARGIN] [%]?

Rationale Recall, for Factors A & B, MSFT had a ***schizophrenic***-profile for: GROSS_MARGIN & EBITDA_MARGIN[%][See Table 7]. Thus, if the Regression Projection where the EBITDA_MARGIN[%] is outside the high side of the 95%CI EBITDA_MARGIN[%] **and** MSFT's GROSS_MARGIN[%] ***stayed inside*** its 95%CI, this would possibly override the concern that an increase in EBITDA_MARGIN[%] would create a decrease in MSFT's GROSS_MARGIN[%] to the extent that it would fall outside the 95%CI[GROSS_MARGIN[%]] on the

Left-Hand Side. This certainly would be a positive expectation or indication for selecting MSFT as Cash-Swap Possibility. Following are these computations made by the DA-Analyst. Using the standard OLS-Regression:

$$\text{GROSS\_MARGIN[\%]} = a + [b \times \text{[EBITDA\_MARGIN[\%]]}$$

we arrive at the following: Note [7]

For EBITDA_MARGIN[%] the EXP[Upper 95%Limit:3.7377] = 42.0012; the X-factor will then be 47.0 [42.0012% + 5%]. The transformation for the Regression will be ln[47.0] = 3.85. The OLS-R projection will be

GROSS_Margin[%] [ln[Estimation]] = 4.3097 = [3.6737 + [0.1652 × 3.85]]

The OLS-R of 4.3097 in the BBT[MSFT:Units] is EXP[4.3097] = **74.4182**

In this case, the GROSS_MARGIN[%] of 74.4 is IN the 95%CI[GROSS_MARGIN[%]] of

EXP[Lower 95%Limit: 4.21935] = 67.9927%→68.0

EXP[Upper 95%Limit: 4.33463] = 76.2944%→76.3

5.5.3 Final Regression Implication

If MSFT, in the future, takes actions to increase EBITDA_MARGIN[%] by 5%, not a trivial increase, the resulting GROSS_MARGIN does not look like it would fall outside of the 95%CI[GROSS_MARGIN] on the low-side; rather the projection would be expect to remain in the population profile of: [68.0 : 76.3]. Our DA-assessment is that this may be a positive indication for the market profile for MSFT if the EBITDA_MARGIN[%] were to increase by 5%. Thus, MSFT is likely, *in this regard*, to be a reasonable Cash-Swap target.

6. Summary & Outlook

6.1 Summary

We have focused on two topics: (i) The ***Nature of the Data*** as the main feature in ***deciding among*** The Harmonic Mean [HM], The Geometric Mean[GM] or the Arithmetic Mean[AM] as the ***data-measure*** for inferential analysis, where: their mathematical order, expected for the trivial case of uniform equality, is: AM> GM> HM, and (ii) The use of the Harman Factor Model [FM] as a screening protocol for ***deconstructing the PPM-correlation network*** so as to garner the intel with respect to a decision-context.

6.1.1 The Nature of the Data

We have presented a discussion and analysis of the nature of the data and its impact on selecting the Data-Measure. The Protocol suggested is very simple, and shockingly, almost exclusively ignored in: (i) academic pedagogic presentations at all levels, (ii) using it as a conditioning measure in data-driven research, such as market-trading studies, and (iii) most textbooks dealing with Data-Analytics. ***Simply, we showed that when there is evidence that the Panel's PPM-association is non-trivial and central tendency-intel is a driver of the inference structure, then using the Arithmetic Mean is convenient but not correct to inform the decision-making process. The clear preference when accuracy and utility of the results are paramount is the Geometric Context***.

6.1.2 Suggested Research Application: Deconstruction of the PPM-Correlation Matrix

Overview The purpose of introducing and discussing a Factor Model [FM] was to indicate *how* an organization, in a research or practical experiential-context, should ***probe the PPM-Network*** of the ***Panels of Data*** selected to inform the decision-making process. The critical phases that were profiled are:

Phase I The DA-Group, the InSource-Advisory, needs to interact with the User-Group interested in investigating a *problématique*, the nature of which requires the analysis of Panels of Data, so as to provide sufficient operational details so that the User-Group can understand. This ultimately ensures that the FM produces the decision-making intel that could aid in addressing their *problématique*,

Phase II The DA-Group offers a Zoom-presentation of the **non-technical-details** of the FM so that the User-Group has a reasonable understanding of how the FM could be used to address the decision-making issues re: the *problématique*.

Phase III The DA-Group & the User-Group will have a joint Zoom-meeting session to arrive at a ***"meeting of the minds"*** as to the process and the benefits of using the FM to address the *problématique*. This will be archived and available to any member of organization,

Phase IV The DA-Group will offer a number of Zoom-meeting presentations of the **technical-details** of ***probing the PPM-Network*** of the ***Panels of Data.*** This will usually require a number of presentation- and Q&A-sessions. After some reasonable period of time there will be a joint-sign-off by the DA- & User- Groups ***to launch the FM-Project***,

Phase V The DA-Group presents possible likely examples drawn from the context of the *problématique* of the User-Group. Of course, ***these practical examples are offered with Q&A-video-links***.

Phase VI Finally, ***during the time working on the project,*** the DA-Group assigns an experienced DA-analyst to be the DA-contact for the User-Group,

Phase VII After the Project has been completed, the HR-Group will collect all the information created **[*Phase I*** through the ***Completion*****]** to present a de-briefing to the DA- & User- Groups with recommendations to ameliorate future DA-InSource Processes.

6.1.3 Typical InSource Deliverables from the MSFT-Case Example

The intent of using the FM is Screening and Identification of the relationships that would likely be of interest in probing the generating processes underlying the creation of market-reported data. To illustrate, the utilization of the FM, we used the Geometric-conditioning as scripted by Carvalho and examined the following seven Variable Panels of the MSFT: {***Four Profile Action-Variables***:{EBITDA, EBITDA_MARGIN[%], EBITA & EBIT} and **Three *Results Measures***: {GROSS_MARGIN[%], OPER_MARGIN[%} & PROF_MARGIN[%]}. We have used the probing protocols to create the following:

I. The Geometric Mean,

II. The Related 95% Confidence Intervals, finally

III. The OLSR-projection for the linear forecasting model:

GROSS_MARGIN[%] = a + [b × [EBITDA_MARGIN[%]]

This set of information, was cast in an assumed Data Analytics Division of a Firm. This offered the context, to script the interactions among: The COO/CFO/CEO, The Finance Department & The Data Analytic Division—a Support Group ["InSource"] to the Firm. Such organizational-configurations are now common fixtures in most MNCs.

6.2 Outlook

Data Analytics[DA] is an exciting, developing and critical "area of interest" that impacts most aspects of the "computational *milieu*". However, it seems to us that that this level of excitation has encouraged Data Analysts to create intel, sometimes before, taking the time to validate the *Nature of the Data* used to create the needed DA-intel. Thus, as our final suggestion, we offer that ALL Data-Analytic-protocols should have: pop-up Excel[VBA™[UserForms©]] of the following nature:

> Generalized *UserForm*:
>
> You have opened the Following Data-Analytic Protocol[DAP] [*Title*]. Before launching this DAP, the following Vetting-Stage questions should be addressed, as our experience suggests that considering these questions and taking the appropriate actions would aid in insuring, insofar as possible, the relevance and utility of the intel that is produced by the Data-Analytic Protocol [*Title*]:
>
> > ¤ **Did you accrue a contiguous set of longitudinally-equally-spaced Panel-Points for the period of interest?**
> >
> > ¤ **Are there at least 13-Panel-Points in the sample?** Note [8] and
> >
> > ¤ **Did you check if there is significant Pearson Product Moment correlation among the Panel-Points.**

The *UserForm* requires a √ in the question's ¤ for each of the vetting-conditions noted above. This *UserForm* will be displayed upon Launch & with the DAP-output. Also, the VBA-protocol will display the User and the Date & Time of Execution. Finally, for this *UserForm* over-rides are not possible—if all the *UserForm* vetting-questions are not completed, the DAP cannot be accessed.

Acknowledgments Thanks and appreciation are due to: Mr. John Conners, Senior Vice President, Financial Counseling, West Coast Region, AYCO for his generous philanthropy which funded the establishment of the John and Diana Conners Finance Trading Lab at the State University of New York College at

Plattsburgh and the Bloomberg Terminals that were instrumental in this research, Prof. Dr. K. Petrova, Department of Economics & Finance, SUNY:Plattsburgh, and the reviewers of the *J. Accounting and Finance Research* for their careful reading, helpful comments, and suggestions.

References

Alvesson, M., Kärreman, D., Sturdy, A. & Handley, K. (2009). Unpacking the client(s): constructions, positions and client–consultant dynamics. *Scandinavian Journal of Management, 25*, 253-263. https://doi.org/10.1016/j.scaman.2009.05.001

Csikszentmihalyi, M. (2000). *Beyond Boredom and Anxiety*. Wiley Editions: [Psychology] pps.272 ISBN: 0787951404, 9780787951405

Czerniawska, F. (2005). The new business consulting landscape. *Consulting to Management* [*Issues & Trends], 16*, 1-6.

Green, P., Tull, D. & Albaum, G. (2009). *Research for Marketing Decisions*, 5th Ed. PHI-Publishing, pps.784. ISBN-9-788120-307575.

Harman, H.H. (1960). *Modern Factor Analysis*. University of Chicago Press. ISBN: 978-1124051710

Kreißig, V. & Taffertshofer, A. (2017). The consequences of client's organizations on management consulting: analysis of consulting projects in Belarus – an experience report. *Journal of East European Management Studies, 22*, 421-438. https://doi.org/10.5771/0949-6181-2017-3-421

Nunnaly, J. (1978). *Psychometric Theory*. 2nd Ed. McGraw-Hill, New York: USA

Rao, C. R., Shi, X., & Wu, Y. (2014). Approximation of the expected value of the harmonic mean and some applications. *Proceedings of the National Academy of Sciences, 111*, 270-273. https://doi.org/10.1073/pnas.1412216111

SAS[JMP6]. (2005). *Statistics and Graphics Guide*. The SAS Institute, Inc., Cary, NC, USA: ISBN-1-59047-816-9

Sturdy, A., Werr, A. & Buono, A.F. (2009). The client in management consultancy research: mapping the territory. *Scandinavian Journal of Management, 25*, 247-252. https://doi.org/10.1016/j.scaman.2009.05.008

Sturdy, A. (2011). Consultancy's consequences? A critical assessment of management consultancy's impact on management. *British Journal of Management, 22*, 517-530. https://doi.org/10.1111/j.1467-8551.2011.00750.x

Tavakil, M. & Dennick, R. (2011). Making sense of Cronbach's alpha. *Int. J. of Medical Education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Notes

Note 1. Prof. Russ Ackoff, The Wharton School of the University of Pennsylvania & Director of the *Anheuser-Busch* Center: Consulting Group of the Social Systems Sciences Dept. of the Wharton School of the University of Pennsylvania.

Note 2. HM, GM & AM discussion-link https://machinelearningmastery.com/arithmetic-geometric-and-harmonic-means-for-machine-learning/

Note 3. In our experience, the In-Source is more than a passing-trend. It is a developing feature that will endure and enjoy organizational permeance. For more background on this exciting development see: Alvesson, Kärreman, Sturdy & Handley (2009), Sturdy, Werr & Buono, (2009), Sturdy (2011) and Kreißig & Taffertshofer (2017).

Note 4. All Zooms are recorded, catalogued by the IT-division, and may be downloaded by firm members.

Note 5. All Factor Models are based upon the work of Prof. Harman. See Harman (1960). We will use as an illustrative example, the Computer-Example, offered by Green, Tull & Albaum [GTA]] (2009), Ch 15[Factor & Cluster Analysis]. The GTA-discussion is highly recommended to capture the essentials of the Harman treatment.

Note 6. In this research report we will be using the Factor Model offered in SAS[JMPv.13]. This is detailed in SAS[JMP6]. (2005) Factor Analysis[450-452]. However, the source of MOST all of the factor model configurations is based upon the work of Prof Harry Harman. His classic text: Harman (1960) is found at: APA: https://psycnet.apa.org/record/1961-02904-000

Note 7. Technical Note: We have used the numbers as reported by Bloomberg. Recall, the Percentages are reported as ##.##. However, if we converted these to decimal-format: 0.#### this would not change the regression values created.

Note 8. Adya & Lusk (2016, p. 74) indicate that in a robust-setting that Panels of the less than 12 Time-Series-Points may compromise the quality of the inferences for most tests.

Appendices

Appendix A Bloomberg[BICS] Industrial Classification System: Selected Firms

AAPL	**AMZN**	**BBBY**	**CMCSA**	**CVS**	**DECK**	**DIS**
EBAY	**FTR**	HSY	**JCP**	KO	**M**	MCD
MSFT	PEP	**RAD**	TAP	**TGT**	TIF	WMT

Table A1 Bloomberg [BICS]-Firms selected for the creation of a DAP. Note 21-Firms were initially selected. Feedback from the Data Analytics Group was that: *Only 13 could be used in the DAP*. The firms Bolded do not fit the DAP: Protocol.

Appendix B Microsoft™, Inc. illustration-Panels

The MSFT[EBITDA_MARGIN[%]] Panel Basic Data from Bloomberg[Panel B1]

39.05	**40.26**	**39.23**	**42.84**	**42.79**	**33.54**	**39.2**
37.97	25.77	36.97	40.49	42.5	44.78	

The LN-Transformation of MSFT[EBITDA_MARGIN[%]] Panel Basic Data from Bloomberg[Panel B2]

3.665	**3.695**	**3.669**	**3.757**	**3.756**	**3.513**	**3.669**
3.637	3.249	3.610	3.701	3.750	3.802	

The MSFT[GROSS_MARGIN[%]] Panel Basic Data from Bloomberg[Panel B3]

79.08	**80.8**	**79.2**	**80.16**	**77.73**	**76.22**	**73.81**
68.82	64.7	64.04	64.52	65.25	65.9	

The LN-Transformation of MSFT[GROSS_MARGIN[%]] Panel Basic Data from Bloomberg[Panel B4]

4.370	**4.391**	**4.371**	**4.384**	**4.353**	**4.333**	**4.301**
4.231	4.169	4.159	4.166	4.178	4.188	

Appendix C Definition of the Variables* used for the DAP[Fin[MSFT{CashSwap}]??

*These are verbatim citations as reported by Bloomberg: Income Statement [GAAP] as of [15June2023]

Excel Field ID: **EBITDA** Measures: Indicator of a company's financial performance which is essentially net income with interest, taxes, depreciation, and amortization added back to it, and can be used to analyze and compare profitability between companies and industries because it eliminates the effects of financing and accounting decisions. Figure is reported in millions; the Scaling Format Override (DY339, SCALING_FORMAT) can be used to change the display units for the field.

Calculated as:

Operating Income + Depreciation & Amortization + Operating Lease Rental Expense Adjustment

Where:

Operating Income or Losses is: IS_OPER_INC,

Depreciation & Amortization is: CF_DEPR_AMORT

Operating Lease Rental Expense Adjustment is: OPERATING_LEA_RENTL_EXPN_ADJUST

Excel Field ID: **EBITDA_MARGIN** [Trailing 12M EBITDA Margin]: Percentage margin of trailing 12-month Earnings Before Interest Taxes Depreciation and Amortization (EBITDA) divided by the trailing 12-month Sales. Unit: Actual.

Calculated as: (Trailing 12-month EBITDA / Trailing 12-month Sales) * 100

Where:

(Trailing 12-month EBITDA / Trailing 12-month Sales) * 100

Trailing 12-Month Sales is TRAIL_12M_NET_SALES

Excel Field ID: **EBITA** Measure calculates earnings before interest, taxes and amortization.

Calculated as: EBITDA - Depreciation Expense

Where:

Depreciation Expense is IS_DEPR_EXP

Excel Field ID: **EBIT** Earnings before interest expenses and income taxes.

Calculated as: Operating Income + Interest Expense

Where:

Operating Income (Losses) is: IS_OPER_INC

Interest Expense is: IS_INT_EXPENSES

Excel Field ID: **GROSS_MARGIN** represents the percent of total sales revenue that the company retains after incurring the direct costs associated with producing the goods and services sold by a company.

Calculated as: (Net Sales - Cost of Goods Sold) * 100 / Net Sales

Where:

Net Sales is: SALES_REV_TURN

Cost of Goods Sold is: IS_COGS_TO_FE_AND_PP_AND_G

Excel Field ID: **PROF_MARGIN** Measuring the company's profitability, this ratio is the comparison of how much of the revenue incurred during the period was retained in income.

Calculated as:

(Net Income / Revenue) * 100

Where:

Net Income is IS050, NET_INCOME

Revenue is IS010, SALES_REV_TURN

Excel Field ID: **OPER_MARGIN** Ratio used to measure a company's pricing strategy and operating efficiency, in percentage.

Calculated as: Operating Income (Losses) / Total Revenue * 100

Where:

Operating Income is: IS_OPER_INC

Total Revenue is: SALES_REV_TURN