K2's The How's And Why's Of Data Analytics

Tommy Stephens

CPA from Woodstock, Georgia

Member, K2 Enterprises

Thirty-eight years of public accounting & private industry experience including twenty-eight years as a CPE discussion leader

BSBA (Accounting), Auburn University and MS (Finance), Georgia State University

Please contact me via email at tommy@k2e.com or Twitter @TommyStephens

Learning Objectives

Upon completing this session, you should be able to:

- List the four types of data analytics
- Identify examples of analytics available in Excel's Analysis ToolPak
- Differentiate between correlation and causation
- Identify at least two analytical techniques available in Power BI

WHAT DATA ANALYTICS AND BUSINESS INTELLIGENCE REALLY ARE

Data Analytics/Business Intelligence

- Many professionals use these two terms interchangeably, and that's understandable because these are related fields
- However, it is important to note that – although related – Data Analytics and Business Intelligence are **two distinct disciplines**, used for **different – yet, often complementary – purposes**
- In modern-day business, both fields have emerged as being **all but necessary in most successful organizations**

What Is Business Intelligence?

Business Intelligence is a set of procedures, tools, and technologies working in concert to transform raw data – potentially vast quantities of raw data – into meaningful information. Managers can use this information to help make decisions and prescribe courses of action that optimize organizational performance.

Why Business Intelligence?

Address the need for more timely and improved decisions

Automate repetitive reports

Identify trends and opportunities

Motivate team members and encouraging teamwork

What Is Data Analytics?

Data analytics is a tool used by BI. More specifically, business professionals use data analytics tools to build data analyses, forecasts, and projections. Data analytics focuses on identifying and managing data that can be useful in identifying ***what happened, why it happened, what will happen in the future, and how we can improve****.*

Data Analytics Answers Four Questions

WHAT HAPPENED?

WHY DID WE ATTAIN A SPECIFIC SET OF RESULTS?

WHAT WILL HAPPEN IN THE FUTURE?

HOW CAN WE ACHIEVE BETTER PERFORMANCE?

Comparing The Two

Data Analytics	Business Intelligence
• At a very high level, **Data Analytics (DA) refers to transforming raw data to make it more useful** • DA attempts to **identify patterns** in data, **relationships between data**, **supports BI**, and often answers the question of "**what will**" happen	• **Business Intelligence** (BI) is typically **used to provide reports** that **help managers with decision-making** processes • Typically, **BI efforts focus on historical data**, seeking to determine **"what" happened** and **"why" it happened**

"BUT I'M AN ACCOUNTANT, NOT A DATA ANALYST"…THINK AGAIN!

We Are All Data Analysts!

- Increasingly, **most of us are being asked to wear the hat of the Data Analyst**…just look at the roles we have assumed
- We are no longer just *forms-fillers*, *paper-pushers*, *bank account- reconcilers* and *ledger-balancers*…technology does that work for us
- And with **Artificial Intelligence (AI)** and **Robotic Process Automation (RPA)** beginning to show up in a big way, that **trend will only continue**!

Accountants Engage In Analytics

- It's **not uncommon for those serving as financial managers to engage in optimization scenarios**
- For example, *"What sales mix would yield the highest gross profit, subject to our company's specific constraints?"*
  - One of the keys to solving this problem is to **identify each significant constraint** in the process
  - Another is to **understand the relationships between the variables**
  - **Once these tasks have been completed, you can begin to construct your forward-looking forecasting and optimization models**

Think About Auditors, For Example

- Many **organizations (not just accounting firms)** are beginning to move down the path of **continuous auditing**, using data analytics tools to alert them in real-time – or very near real-time – of anomalies in the data they are observing
- This potentially **reduces the need to sample data** and provides the **opportunity to examine 100% of the population in real-time** or very near real time
- As an example, consider the **opportunity for auditors to incorporate social media statistics into projections for future sales for a client that operates in the retail space**

A VERY BIG HURDLE TO OVERCOME: *UNDERSTANDING CLIENT NEEDS*

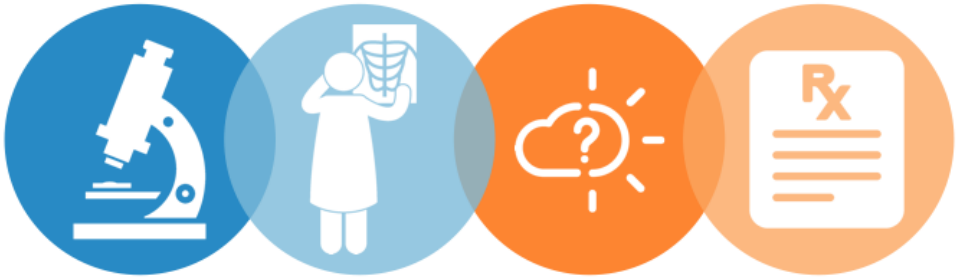
An Example…

- Someone may ask the question "*what is driving sales*"
- The **Accountant** might likely rush to their computer and generate various sales reports – *by customer*, *by item*, *by region*, *by salesperson*, *by channel*, *by location*, *etc*.– but this provides *descriptive* information…and, unfortunately, it doesn't answer "*what is driving sales"*
- Don't be misled, **there is value in these reports, but it doesn't really help to answer the question "what is driving sales"**
- The **Data Analyst** would look to other data…

An Example…

- Once the Data Analyst identifies all the relevant drivers of sales, they would begin to *identify the relationships in the data*, *create a data model*, *regress the data*, and *use the outputs to explain what drove sales* ***in the past and, perhaps, build a model***
- However, **while the work outlined above may be useful in predicting the future, there is no guarantee that such a prediction will be useful** or meaningful
- Why not? It's simple: **conditions and circumstances change** and as they do, we must revise our models, methods, and formulas

Four Types Of Data Analytics



Descriptive
Explains what happened.

Diagnostic
Explains why it happened.

Predictive
Forecasts what might happen.

Prescriptive
Recommends an action based on the forecast.

Four Types Of Data Analytics

- A report comparing total sales over the prior five years to total credits issued would be ***descriptive analytics***
- An auditor might use regression analysis to forecast, for example, future credit memos past on historical data…***predictive analytics***
- The auditor might also use diagnostic analytics to determine why significant variances between the forecast and actual results materialized…this would be ***diagnostic analytics***
- ***Prescriptive analytics*** might be used to create an action plan for bringing credit memos back in line with expectations

Descriptive Analytics

- **Descriptive analytics** addresses “what happened”
- Often includes historical data and benchmarks
- Helps to identify strengths and weaknesses, where targets/budgets were met and where they were missed
- Examples, month-over-month sales, actual results compared to budget, status of KPIs, etc.
- Oftentimes found in historical accounting data, but **not limited to accounting data**

Diagnostic Analytics

- The analyst uses **diagnostic analytics** to answer “why”
  - *Why did we exceed our targeted sales number on a product?*
  - *Why are our customer satisfaction scores below target?*
  - *Why are we exceeding our budget for overtime expenditures?*
- Diagnostic analytics helps to perform such actions as **identifying outliers**, **uncover previously-hidden relationships** in the data, and **isolate various patterns and trends**

Predictive Analytics

- Use **predictive analytics** when attempting to predict the future
  - *Given a $1 million advertising expenditure, what will incremental sales be for the next quarter?*
  - *At this burn rate, how much cash we will have left at the end of the third quarter?*
  - *Are we on target to meet our annual sales goal?*

Prescriptive Analytics

- Use **prescriptive analytics** to predict the future
  - For example, *what level of advertising expenses should we commit to achieve our sales goal*?
  - Lead scoring…*if someone makes repeated visits to our web site, aren't they likely to buy from us in the future*
  - *Analyzing bank transactions for potential instances of fraud by comparing current transactions to historical norms*
- **Machine learning** and **artificial intelligence** sometimes plays a role in these activities

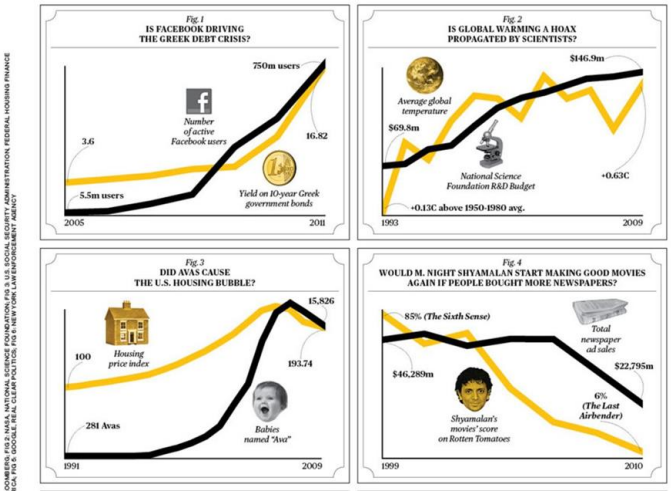
Correlation ≠ Causation

- **Correlation DOES NOT Mean Causation!**
- **Causation** is an occurrence or action that can cause another
- **Correlation** is an action or occurrence that has a direct link to another
- In **causation, the results are predictable and certain** while **in correlation, the results are not visible or certain but there is a possibility that something will happen**

A Very Big Caveat!

- **Correlation DOES NOT Mean Causation!**
- For example, **regression analysis** examines the relationship(s) between a dependent variable and an independent variable(s)
- Often, **regression analysis can be used to identify and define cause-and-effect relationships** and help us to **predict the future with reasonable certainty**
- However, **just because two (or more) variables are correlated, you CANNOT infer a cause-and-effect relationship**
- Consider the case of **thunderstorms relative to ice cream sales**

Correlation & Causation



Some examples of ***correlation, yes,*** *but* ***causation, probably not***

Common Data Analytics Tools

Everyday Data Analytics Tools

Excel	Excel Add-ins
Power BI Desktop	Power BI

EXAMPLES OF DATA ANALYTICS USING MICROSOFT OFFICE EXCEL

Descriptive Analytics Using Excel

- Although you can use various Excel functions – **AVERAGE**, **AVERAGEIF**, **STDEV**, **MAX**, **MIN**, and others to generate to descriptive analytics in Excel, perhaps there is no easier way to do it than by using the **Analysis ToolPak**
- However, **it is an add-in provided by Microsoft** and you must **activate it before you can use it**
- Once you activate it, you can use its **Descriptive Statistics feature to generate at least 13 different descriptive statistics**

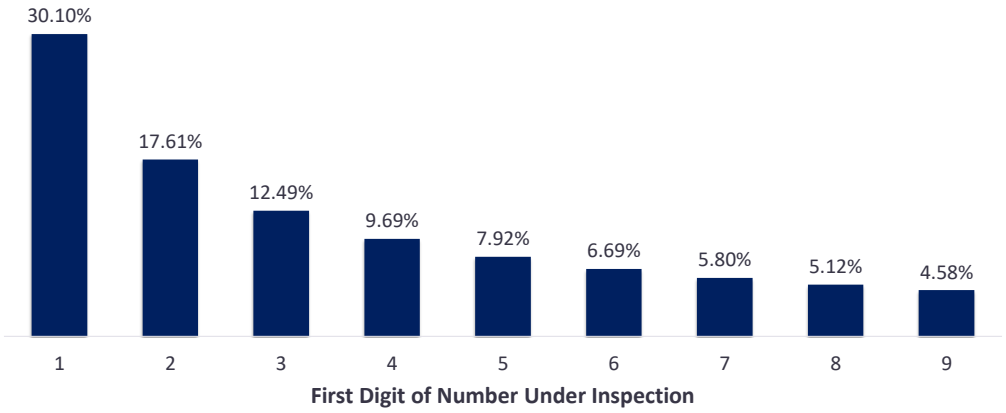
Descriptive Analytics Using Excel

- You can also use the Analysis ToolPak to generate **Histograms**, another form of descriptive statistics
- **Histograms group data into various "bins,"** helping to potentially identify concentrations of risk
- The advantage of using the Analysis ToolPak for this application (and others) is simple…**you don't have to create the formulas**!

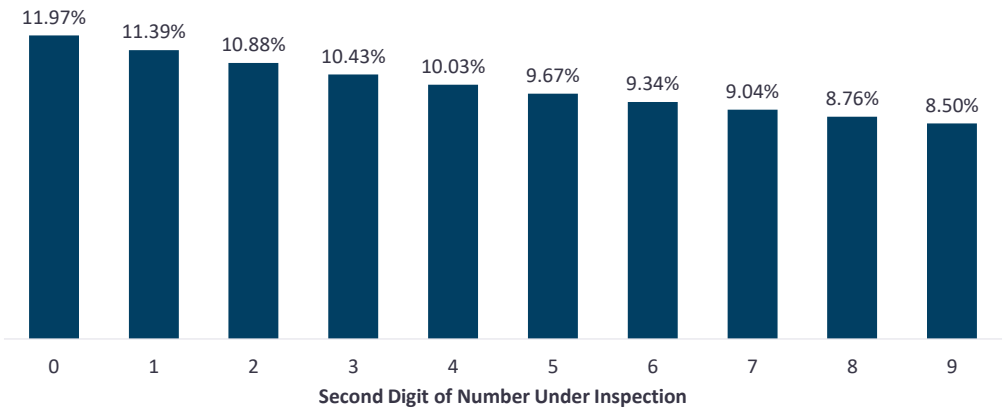
Descriptive Analytics Using Excel

- You can also use histograms to test data for comporting to **Benford's Law**
- In sum, Benford's Law says that numbers do not follow a random walk; rather they follow an established pattern
- If significant deviations exist from the pattern, you may want to examine for errors or fraud

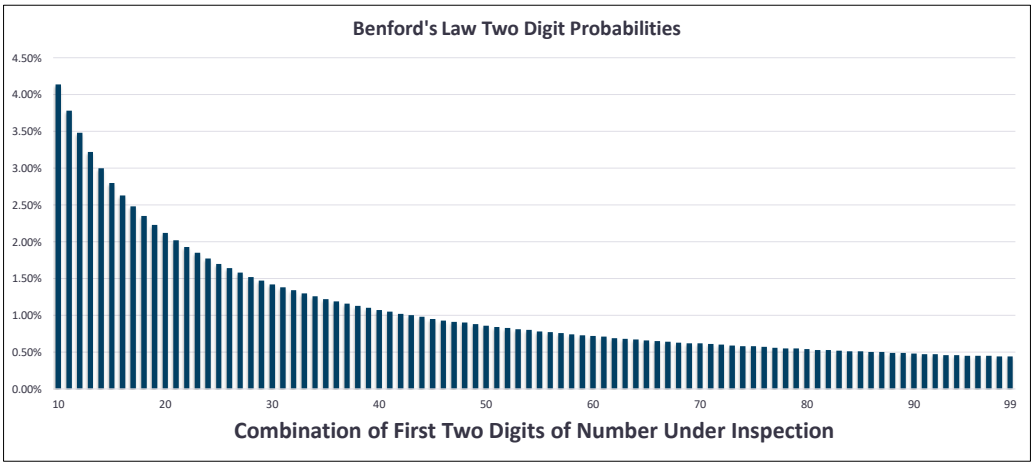
Benford's First Digit Probabilities



Benford's Second Digit Probabilities

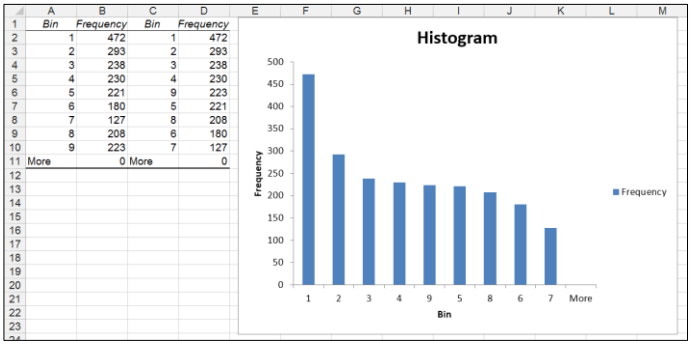


Benford's Two Digit Probabilities



Benford's Law

- You can use the Analysis ToolPak's **Histogram** tool to create the groupings for a Benford's test by extracting the first digit from each number in a series
  - Use Excel's **LEFT** function for this step
- Then, create a Histogram with "1" through "9" as the various bin groupings



	A	B	C	D
1	*Bin*	*Frequency*	*Bin*	*Frequency*
2	1	472	1	472
3	2	293	2	293
4	3	238	3	238
5	4	230	4	230
6	5	221	9	223
7	6	180	5	221
8	7	127	8	208
9	8	208	6	180
10	9	223	7	127
11	More	0	More	0

Limitations Of Using Benford's Law

- First, **large sets of data are more likely to fit Benford's Law** than smaller sets
  - As with virtually all mathematical models, large sets of data are more reliable
- Second, **any set of data based on assigned numbers will likely not follow Benford's Law**
  - A good example would be Social Security Numbers, which are issued based on a prescribed order and methodology
- **Anomalies do not necessarily mean fraud or errors**…they could be just simple aberrations in the data

Excel's Analyze Data Feature

- Excel's **Analyze Data** feature (formerly known as Ideas) can take **a large data set and quickly summarize and analyze it**
- The output can help you to **identify trends and relationships** in your data that you may not have known existed
- Additionally, **you can "ask questions" about your data by just typing them in Excel**

Predictive Analytics Using Excel

- As with descriptive analytics, **numerous options exist for generating predictive analytics**
- The most common methods include Excel's **FORECAST** and **FORECAST.ETS** functions and the **Forecast Sheets** feature
- **These tools can predict the future based on the relationship between a dependent variable and an independent variable**
- **FORECAST.ETS identifies seasonality** in the data and incorporates it into the forecast and **Forecast Sheets allows for confidence intervals**

Regression Analysis In Excel

- For more advanced predictive analytics, you could use **Excel's Regression Analysis tool**, found in the Analysis ToolPak
- You can build your forecasts on up to **sixteen** independent variables
- The **tool will generate an R-square value and a correlation coefficient,** both of which will help you to understand the "fit" of the data
  - **Correlation coefficients greater than 0.8 or less than -0.8 generally indicate a good "fit" of the data**
- Remember that discussion on correlation and causation? **It's important here!**

An Excel-Based Diagnostic Example

- Not only can **regression analysis** be a good predictive analytics tool, it **can also assist with diagnostic analytics**
- For example, with regression analysis, we may be able to identify (based on higher R-square values) which variables have the greatest impact on whether we achieve (or miss) an intended result
- We also may be able to use the **Analyze Data** feature to find "hidden" relationships in our data

Solver For Prescriptive Analytics

- You can use **Solver** (another Excel add-in) **to assist in identifying the proper mix of inputs to achieve the best result**
- Solver models can identify the inputs necessary to realize **maximum values**, **minimum values**, or **specified values** – subject to **user-defined constraints**
- Note, this is **NOT** the same as **Goal Seek**

USING POWER BI FOR DATA ANALYTICS

Analytics Options With Power BI

- Power BI can generate all four types of analytics
  - Descriptive analytics calculating measures, including **ten different means of summarizing data**
  - Diagnostic analytics using the **Analyze** feature
  - Predictive analytics by **generating forecasts from line charts, complete with confidence intervals**
  - Prescriptive analytics using **Dataflow** features

THANK YOU!