



Practical Ways to Utilize Statistics

Jacob Rouse

Presentation Description

Examples on how we can dust off our old statistic books and understand our operations.

Is the process in statistical control?

This presentation will walk you through examples of determining actions that should be taken based on statistical performance.

Ever wonder if there is a correlation between Packer's results and absenteeism? Come check out the answer and how you can answer similar questions

Please note: all sources are included in presentation "Notes" section

Session Objectives

Understand a number of key statistic concepts that have practical use

Review how these statistical concepts can be easily utilized in Excel

Provide a number of examples how various operational areas can utilize statistics immediately



Introduction: Jacob Rouse (Presenter)

Introduction: Jacob Rouse (Presenter)

Proud husband and father of four young kids. Health, improvement, family, community, and faith are the values that drive my work, community involvement, personal lifestyle, and decision to work for a Population Health Organization.

- *Chair of local United Way Investment Counsel, WI HFMA board member, finance and advisory committee member of church and other local non-profit organizations.*
- *Heavy involvement as soccer coach in community.*
- *Avid DIYer*
- *Experience: revenue cycle, rolling budgeting, forecasting, self-pay account receivable process (statement redesign, payment plans, utilization of online platform), retail operations, athletic training*
- *Director of Revenue Cycle (supporting Belin's front to back cycle departments and functions)*
- Contact Info
 - Jacob.Rouse@bellin.org
 - Office: 920-445-7211
 - **Cell: 763-227-0651**



Refresh



Refresh

- What are “statistics”?
 - **Simple Language: Math using data to understand the world. Helps prove or disprove assumptions.**
 - Technical Language: the practice or science of collecting and analyzing numerical data in large quantities, especially for the purpose of inferring proportions in a whole from those in a representative sample.



Statistical Control



Terms

- **Statistical process control (SPC)**

- **Simple Language: statistics that help you understand process results**

- *Technical Language: Statistical techniques to control a process or production method. SPC tools and procedures can help you monitor process behavior, discover issues in internal systems, and find solutions for production issues.*

- **Statistically significant**

- **Simple Language: test to quantify if result was by chance or influenced by something**

- *Technical Language: Determination about the null hypothesis, which hypothesizes that the results are due to chance alone. A data set provides statistical significance when the p-value is sufficiently small*

- **Standard Deviation**

- **Simple Language: measure of how different data points in a data set are**

- *Technical Language: a measure of the amount of variation or dispersion of a set of values. A low standard deviation indicates that the values tend to be close to the mean (also called the expected value) of the set, while a high standard deviation indicates that the values are spread out over a wider range.*

Real Life Example

Question: is my monthly fluctuation in AR days “appropriate”?

How to statistically answer:

- **Step 1: Excel sheet with data**
 - 3 years data by month with AR Days
- **Step 2: Formulas (excel formulas in notes)**
 - Average – (find the “mean”)
 - Standard Deviation (what is “normal” variation?)
- **Step 3: Add Columns and create chart**
 - **Column C = Mean**
 - **Column D = Mean + 1 STDEV**
 - **Column E = Mean – 1 STDEV**

Mean	45.30
Standard Deviation	3.38

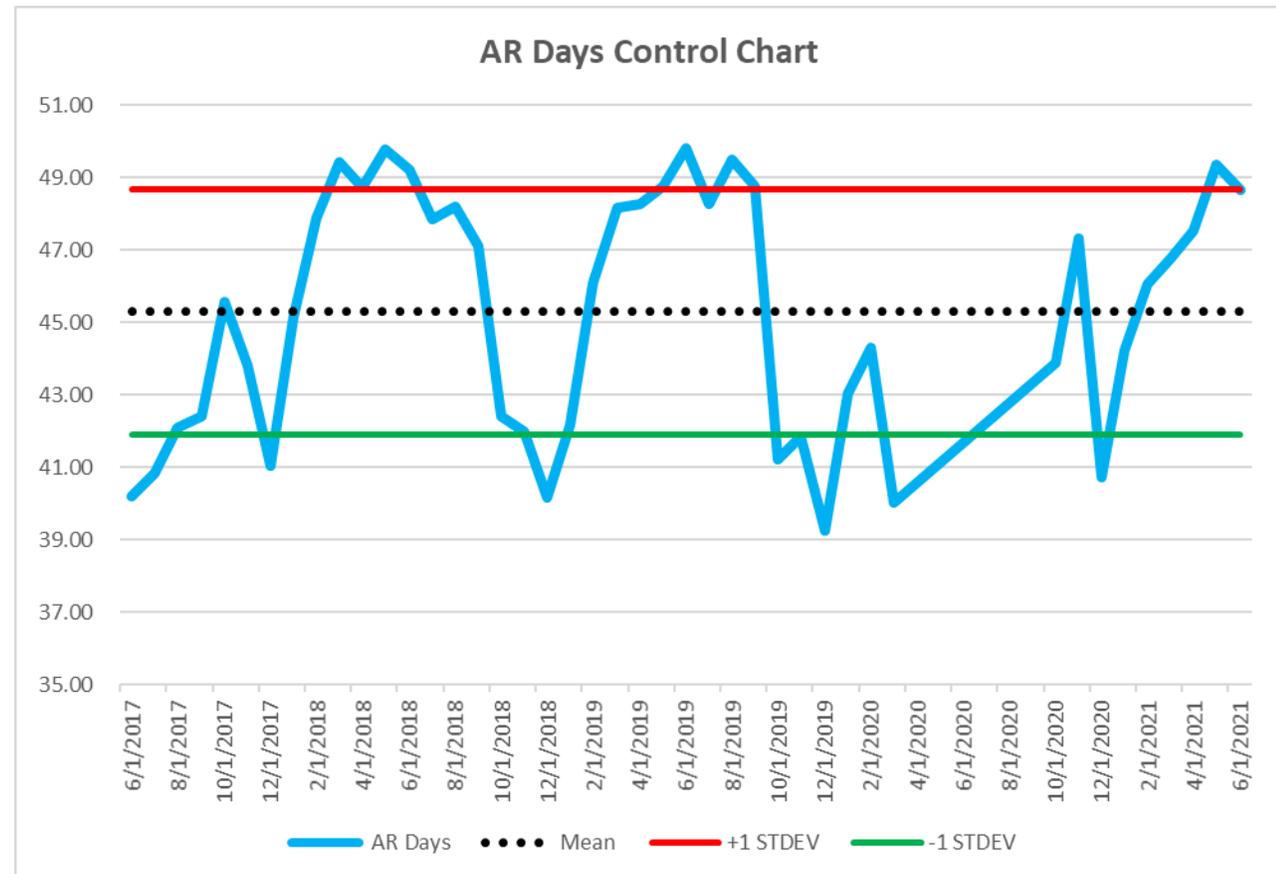
Initial Data		Added Data		
Month Year	AR Days	Mean	+1 STDEV	-1 STDEV
6/1/2021	48.66	45.30	48.68	41.92
5/1/2021	49.36	45.30	48.68	41.92
4/1/2021	47.55	45.30	48.68	41.92
3/1/2021	46.76	45.30	48.68	41.92
2/1/2021	46.06	45.30	48.68	41.92
1/1/2021	44.25	45.30	48.68	41.92
12/1/2020	40.72	45.30	48.68	41.92
11/1/2020	47.31	45.30	48.68	41.92
10/1/2020	43.88	45.30	48.68	41.92
3/1/2020	40.04	45.30	48.68	41.92
2/1/2020	44.32	45.30	48.68	41.92
1/1/2020	43.03	45.30	48.68	41.92
12/1/2019	39.26	45.30	48.68	41.92
11/1/2019	41.86	45.30	48.68	41.92
10/1/2019	41.23	45.30	48.68	41.92
9/1/2019	48.75	45.30	48.68	41.92
8/1/2019	49.49	45.30	48.68	41.92
7/1/2019	48.28	45.30	48.68	41.92
6/1/2019	49.80	45.30	48.68	41.92
5/1/2019	48.76	45.30	48.68	41.92
4/1/2019	48.28	45.30	48.68	41.92
3/1/2019	48.16	45.30	48.68	41.92
2/1/2019	46.10	45.30	48.68	41.92
1/1/2019	42.17	45.30	48.68	41.92
12/1/2018	40.17	45.30	48.68	41.92
11/1/2018	42.01	45.30	48.68	41.92
10/1/2018	42.42	45.30	48.68	41.92
9/1/2018	47.12	45.30	48.68	41.92
8/1/2018	48.20	45.30	48.68	41.92
7/1/2018	47.86	45.30	48.68	41.92
6/1/2018	49.21	45.30	48.68	41.92
5/1/2018	49.78	45.30	48.68	41.92
4/1/2018	48.72	45.30	48.68	41.92
3/1/2018	49.44	45.30	48.68	41.92
2/1/2018	47.88	45.30	48.68	41.92
1/1/2018	45.17	45.30	48.68	41.92
12/1/2017	41.05	45.30	48.68	41.92
11/1/2017	43.81	45.30	48.68	41.92
10/1/2017	45.56	45.30	48.68	41.92
9/1/2017	42.43	45.30	48.68	41.92
8/1/2017	42.08	45.30	48.68	41.92
7/1/2017	40.85	45.30	48.68	41.92
6/1/2017	40.20	45.30	48.68	41.92

Real Life Example

Question: is my monthly fluctuation in AR days “appropriate”?

Answer / Summary:

- Process is designed to stay in-between red and green lines
- Anything “outside” is abnormal
- **What to do**
 - **Outside** = create plan to get inside or re-establish baseline
 - **Inside** = Accept result
 - **Don't like results?** = Process needs redesign (see next slide)





Process Result Overview

“Every system is perfectly designed to get the results it gets” – Edward Deming

Deming Bio

- American engineer, statistician, professor, author, lecturer, and management consultant
- Leading management thinker in the field of quality

Takeaway of analysis and quote:

- Results won't change unless process changes
- [ASQ](#) offers tools to plan and redesign





Correlation



Terms

- **Correlation**

- **Simple Language: there a direct relationship between variables**

- *Technical Language: Correlation is a statistical measure that expresses the extent to which two variables are linearly related (meaning they change together at a constant rate). It's a common tool for describing simple relationships without making a statement about cause and effect.*

- **F-Test**

- **Simple Language: statistical test to prove 2 samples can be compared (checks variation)**

- *Technical Language: F-test is any statistical test in which the test statistic has an F-distribution under the null hypothesis. It is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled*

- **T-Test**

- **Simple Language: statistical exercise to prove if 2 data sets are the same**

- *Technical Language: A type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features.*

- **P-Value**

- **Simple Language: the key output of a T-Test to answer if 2 samples are the same**

- *Technical Language: The probability of obtaining results at least as extreme as the observed results of a statistical hypothesis test, assuming that the null hypothesis is correct. If F-Test shows there is a*

Real Life Example

Question: do the Packer's results impact absenteeism?

Is there a higher ratio of unplanned time off to scheduled worked hours the day after?

- **Samples comparing:**

- **No Game Mondays** = Mondays with no Packers games on Sunday or Monday
- **Mondays After Game** = Mondays after Sunday Packer Game
- **Mondays After Win** = Mondays after Sunday Packer Game with Win
- **Mondays After Loss** = Mondays after Sunday Packer Game with Loss
- *Also reviewed: entire population (all days), Sundays without games, each weekday, game days)*

- All comparing samples have F-Test under .02 (meaning: all statistically appropriate to compare)



Real Life Example

Question: do the Packer's results impact absenteeism?

Is there a higher ratio of unplanned time off to scheduled worked hours the day after?

How to statistically answer:

- **Step 1: Excel sheet with data**

- 3 years data by day of unplanned and scheduled hours with Packer results identified

- **Step 2: Formulas**

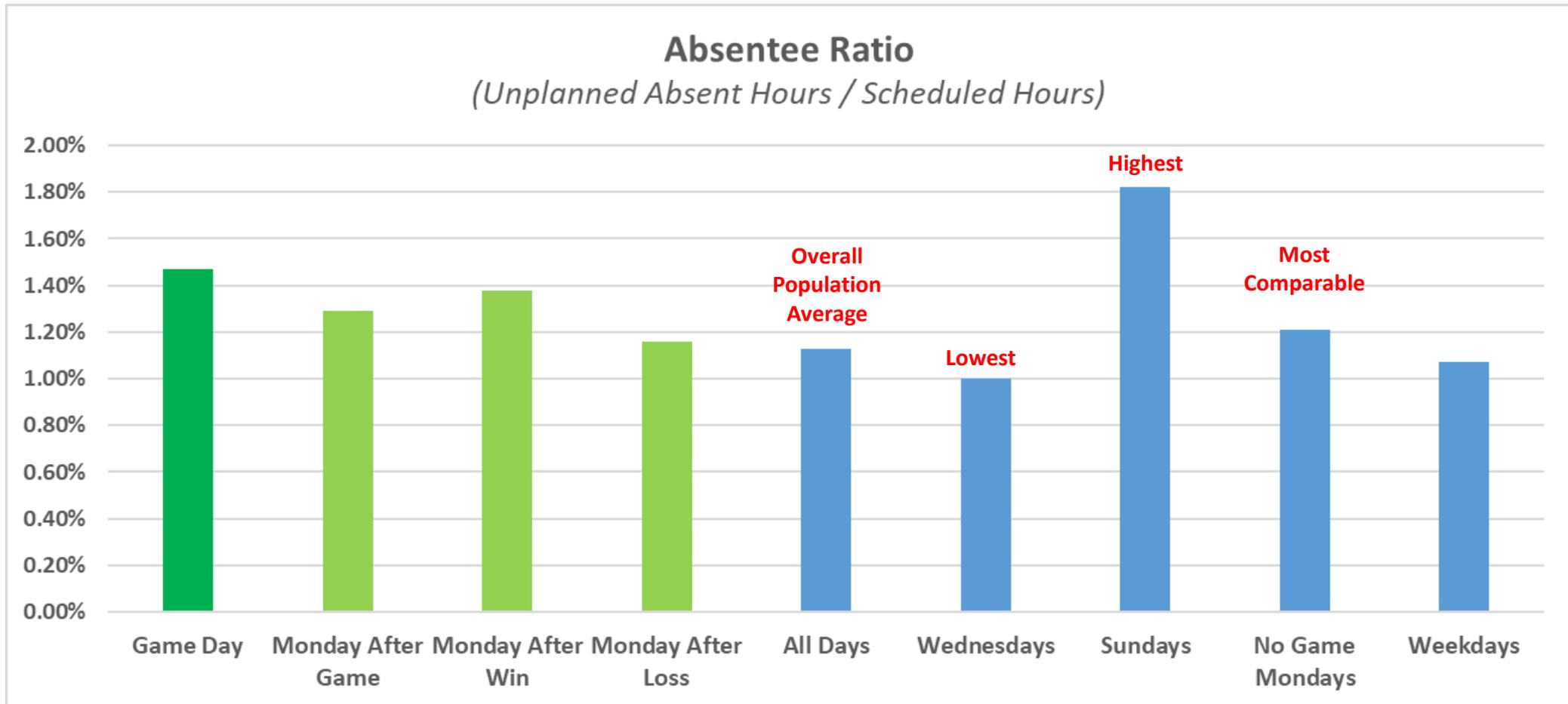
- Ratios (Unplanned / Scheduled during different periods)
- Step 1 = F-Test & Step 2 = T-Test (Google and YouTube have many walkthroughs and notes contain summary steps)

- **Step 3: Analyze results**

- See next slides

Real Life Example

Question: do the Packer's results impact absenteeism?



Real Life Example

Question: do the Packer's results impact absenteeism?

Key Outputs of T-Test	Meaning
Mean	Average of samples
Observations	Number of data points
df	Sum of both samples data points
P(T <= t) two-tail	P-Value (.05 = statistically significant)

No Game Monday vs Day After Sunday Game

No Monday Game	Monday After Game
1.26%	1.27%
150	64
212	
0.951	

No Game Monday vs Day After Sunday Game WIN

No Monday Game	Monday After Win
1.26%	1.37%
150	38
186	
0.389	

No Game Monday vs Day After Sunday Game LOSS

No Monday Game	Monday After Loss
1.26%	1.13%
150	25
173	
0.357	

Monday after Sunday Win vs Monday after Sunday Loss

Win	Loss
1.37%	1.13%
38	25
61	
0.089	

Statistical test = "t-Test: Two-Sample Assuming Equal Variances"

- Simple Language: tests if the samples are statistically the same

Results: There is a clear difference in outcomes (.05 and under = statistically same)

Tells Us: 95.1% change that that results are NOT the same.

Results: There is a clear difference in outcomes (.05 and under = statistically same)

Tells Us: 61.1% chance Win vs No Game will have same absenteeism

Results: there is a clear difference in outcomes (.05 and under = statistically same)

Tells Us: 64.3% chance Loss vs No Game will have same absenteeism

Result: Closest P-Value (almost statistically significant).

Tells us: there is almost no statistical difference in absenteeism Win vs Loss

Real Life Example



Question: do the Packer's results impact absenteeism?

Results Summary

- **Sunday Packer games DO lead to higher absenteeism on Monday**
- **Though Mondays after wins have higher absenteeism than those with losses, there is strong evidence there is NOT a direct correlation.**
- **Sundays with Packer games have lower absenteeism than Sundays without Packer games and this IS statistically significant.**

Real Life Example

Question: do the Packer's results impact absenteeism?

Results Detail

- Absenteeism is **19%** greater when Packers Win vs Lose **BUT** statistically, this is nearly impossible to prove the correlation (**P-Value .08**)
- Mondays without Sunday Packer Games are only **6%** less likely to have absenteeism than those with games day before **BUT** statistically, there is a clear difference in the correlation (P-Value .95)
- Sundays with Packer games have **19%** less absentee hours than those without **AND** there is a statistical difference in results (**P-Value = .35**)

Reminder

- Correlation Does Not Imply Causation MEANING **the statistical outcomes show correlation but not necessarily reason**

Ideas on How YOU can use

- ER visits on full moon vs non-full moon
- Admission acuity by day
- COVID reported infection to COVID hospitalization timing

Questions?