

**School of Informatics & IT (IIT)**

**School of Applied Science (ASC)**

**AY2024/2025 Oct Semester**

**IDL (Inter-disciplinary Learning) Assignment**

**CDA2C04 Data Storytelling**

**&**

**APH3021 Pharmaceutics and Compounding**

Submitted by: Group 16

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Student Name** | **Student Admin No.** | **Student Diploma** | **Student Class** |
| 1 | Adam Haizad Bin Mohamad Faizal | 2302276G | Big Data & Analytics | PC02 |
| 2 | Iffa Umairah Binte Mohamed Ismail | 2301953E | Pharm Sci | TAL2 |
| 3 | Iman Abdul Ra’uf | 2300811C | Pharm Sci | TAL2 |
| 4 | Muhammad Muaz Bin Azmee | 2305306I | Big Data & Analytics | PC02 |

**Declaration of Originality**

I am the originator of this work and I have appropriately acknowledged all other original sources used as my references for this work.

I understand that Plagiarism is the act of taking and using the whole or any part of another person’s work, including work generated by AI, and presenting it as my own.

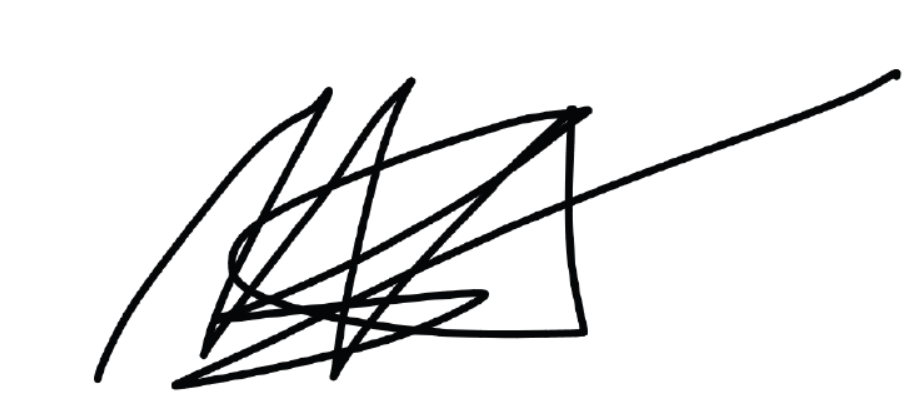
I understand that Plagiarism is an academic offence

and if I am found to have committed or abetted the offence of plagiarism in relation to this submitted work, disciplinary action will be enforced.



Iman

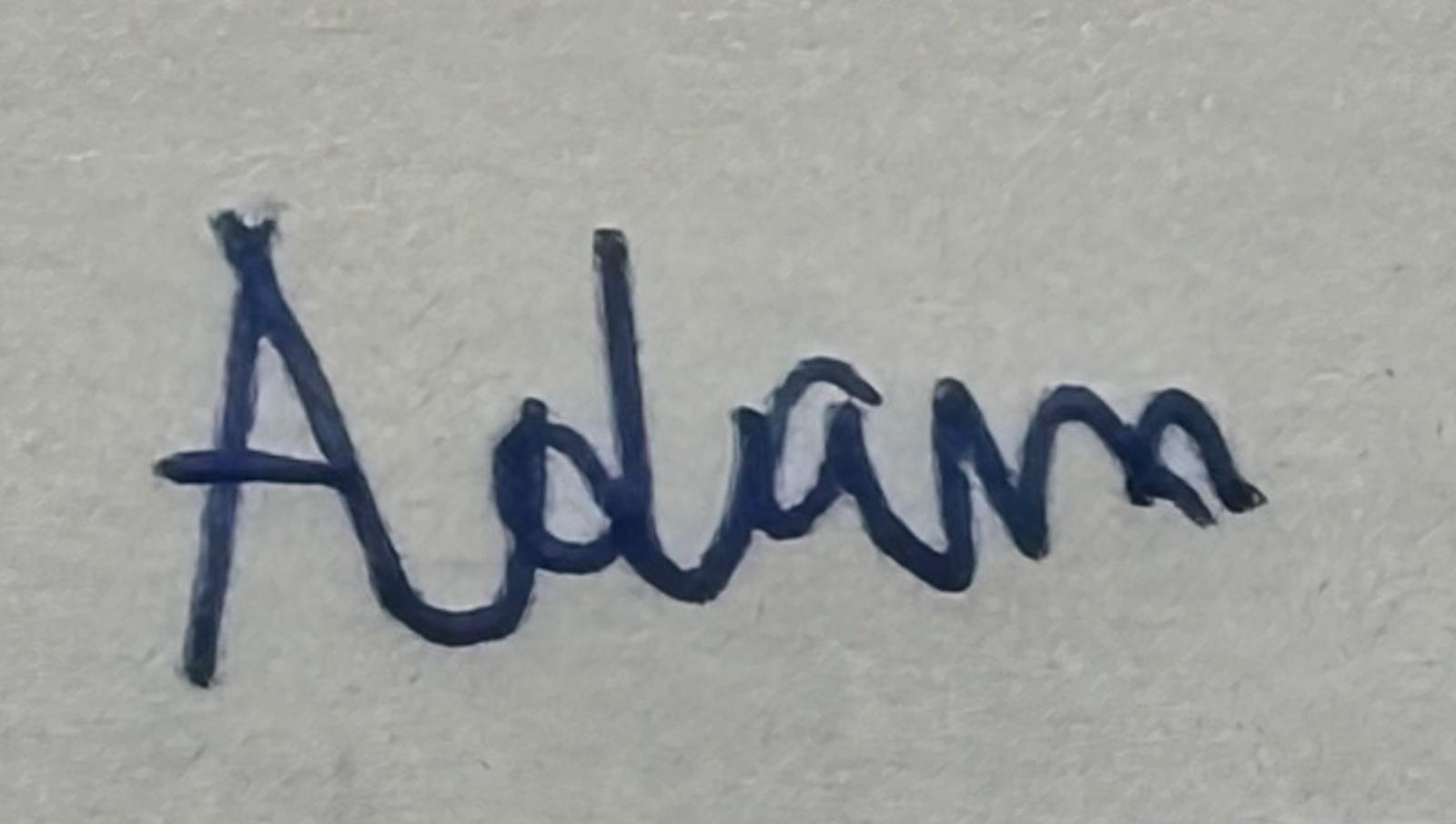
**Name and Signature of <student>: ……………………………………**



**Name and Signature of <student>: ……………Muaz………………**



**Name and Signature of <student>: ………Iffa ……………………**



**Name and Signature of <student>: ……Adam ……………………**

**Name and Signature of <student>: ……………………………………**

**Date <signing date in dd /mm/yyyy format>:** **…11/2/2025…………………………**

# **Declaration on the use of Generative AI tools for assignments**

|  |
| --- |
| Describe how you have used Generative AI tools such as ChatGPT or Dall.E- in your assignment.    Share the link to the conversations you had with the AI tool (i.e., the prompts you used and the responses you get from the AI tool).    **Please refer to this PDF on “How to share the conversations made with ChatGPT?”** |
|  |
| How do you indicate the reference?  The content generated by AI tools are not retrievable except by the user who generated them, so they are considered non-recoverable sources. Although non-recoverable data or quotations in APA Style papers are usually cited as personal communications, with ChatGPT-generated text there is no person communicating. Quoting text from ChatGPT chat is therefore more like sharing the output of an algorithm, with a reference list entry and the corresponding in-text citation.  According to the official APA Style site, ChatGPT references should be cited as:  E.g. OpenAI. (2024). *ChatGPT* (Feb 13 version) [Large language model].  <https://chat.openai.com/chat>    Note: The information in parentheses refers to the update or revision date of the model used. Refer to the release notes in the ChatGPT application. |

**Important Note:**

· Do not copy answers produced by the AI tool in totality as it is considered as plagiarism.

· Do not rely on any information produced by the AI tool blindly. You should always verify the answer with other sources. Do not assume that these answers provided by the AI tool are correct.

· To achieve quality outputs from the AI tool, you should provide good prompt that is clear and specific. Be precise and provide context. Avoid asking open-ended questions.

# 

# **1.** **Business Understanding**

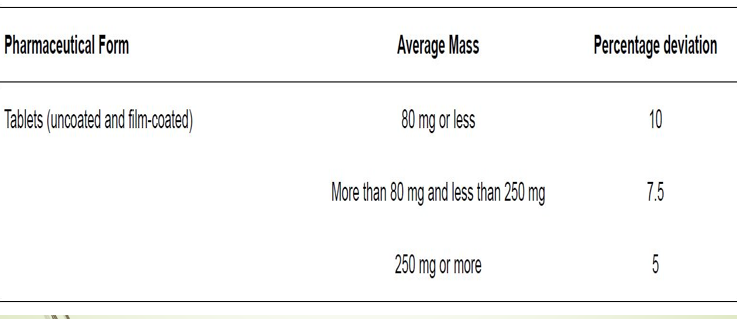
1. **Determine the target audience for your analysis:**

* Various regulatory bodies such as
  + QA (quality assurance) and QC (quality control) teams
  + Senior management
  + External auditors or inspectors
  + stakeholders

1. **Understand the Pharmacopoeia Standards Applicable to your dataset**

Requirement to pass Uniformity of Weight (Mass) according to british pharmacopoeia (BP):

* Not more than 2 of the individual masses deviate from the average mass by more than the percentage deviation shown in Table 2.9.5.-1 and
* None deviates by more than twice that percentage.



1. **Determine the method for testing the Uniformity of weight (mass) of the tablets**

* Every 15 minutes, 20 tablets are made by the machine and sampled
* For each time frame, for example at the 15 minute mark, all 20 tablets are used to Calculate avg mass for each sample
* Find percentage deviation of single tablet mass from the average using formula : (|average - mass| / average) x 100%
* Note: Use absolute | | to avoid a negative percentage
* Using the dataset given, each tablet weighs below 80mg. The acceptance criteria would be the percentage deviation should not be more than 10%.
* If the percentage deviation is more than 10%, it means that that batch of 20 tablets have failed the uniformity of weight test → this means that the production will have to be halted until the underlying issue has been rectified. Refer to part d for the possible reasons to why the test had failed.
* If percentage deviation is less than 10%, it means that it has passed the uniformity of weight test, making the tablets suitable for production under those fixed conductions that they were produced under.

1. **Identify Key Factors Influencing Uniformity of Weight (Mass) in tablet production**

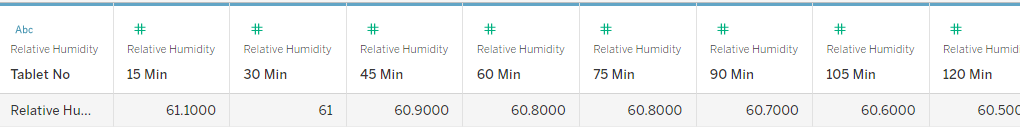
* Two most important properties the uniformity of weight (mass) in tablet production:
* Powder flow: fluidity of powder from hopper to machine
  + Increase powder flow by adding glidant
  + Make particles as spherical as possible
  + By granulation
  + By vibrators
  + Some factors involved include:
    - Particle size distribution
    - Bulk & tapped density
    - Compressibility index
    - Hausner ratio
* Compressibility
  + A property of forming a stable and intact compact mass when pressure is applied
  + Materials that do not compress produce soft tablets
  + Granulation improves compressibility
* Other factors that may influence the uniformity of weight in tablet production:
* Environmental conditions
* Tablet size and shape
* Filling accuracy
* Active Pharmaceutical Ingredient (API) properties

# **2.** **Data Management**

1. **Perform data audit to identify any data quality issues. This is to check the dataset for completeness and validity.**

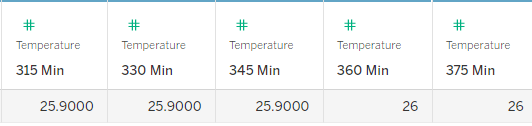
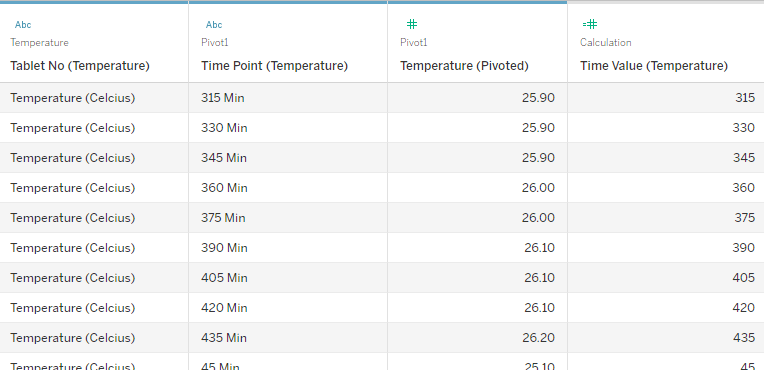
Relative Humidity

Before After

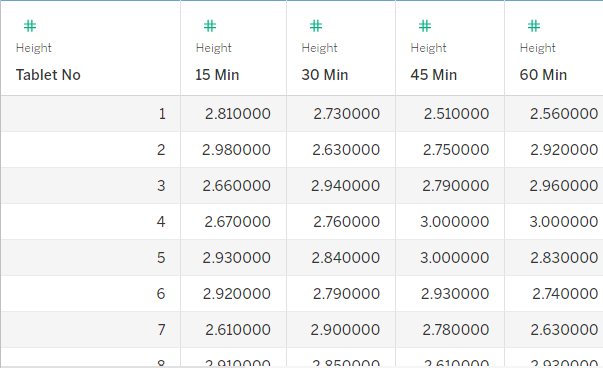
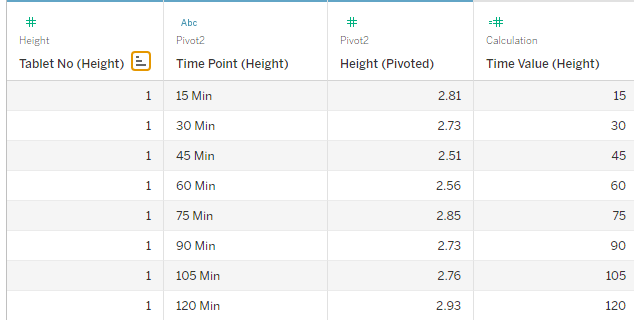


Temperature

Before After

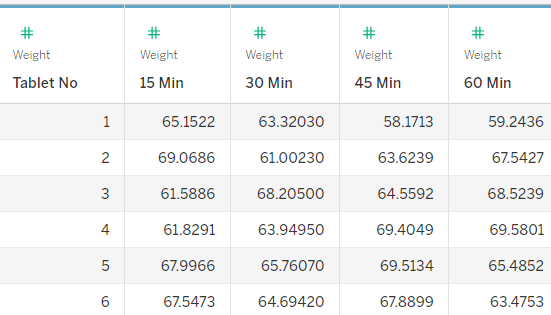


Height

Before After

At first, the values for the Relative Humidity, Temperature and Height decimal precision varied from row to row, with some values showing up to four decimal places and others showing none. To ensure standardization, We first pivoted the table into a tall table format for easier conversion (further explained in (b)). We then used Tableau's Number style function to standardise all numeric fields to two decimal places. This ensured consistency, enhanced readability and removed any possible problems brought on by inconsistent accuracy.

1. **Evaluate how you may want to reshape your data to the appropriate structure (e.g. tall table) for analysis purposes**

Before After

At first, the dataset was in a wide format with time points in distinct columns, which made it challenging to sort by time or analyse trends. We converted the data into a tall format using Tableau's Pivot function, merging all time columns into a single "Time Point" column with matching values. We made a Time Value field to extract the time's numeric component for accurate sorting because "Weight (Pivoted)" stored time as strings (for example, "15 Min"). This change simplified the dataset for calculations and visualisations and allowed for appropriate chronological analysis.

1. **Based on the pharmacopoeia standards, compute all parameters and explain in detail the method(s) used and computation(s) done for the parameters.**

The question calls for using Tableau to calculate pharmacopoeia parameters. Using pharmacopoeia formulas, important parameters such as average mass and percentage deviation are computed.

The formula are as followed:

Average Mass = Sum of tablet weights at a time point / Number of tablets

Percentage Deviation = (|average mass - tablet weight| / average mass) x 100%

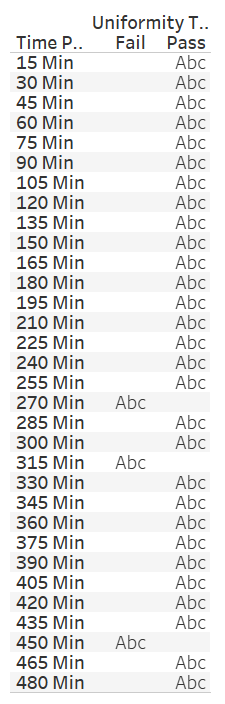
We then created a new calculated field for the Average Mass:  


The FIXED calculation ensures that, regardless of other filters, the average is computed within the group of each time point.

The same is done for the Percentage Deviation: 

1. **Calculate the average mass of tablets at each time point, identify the corresponding percentage deviation, and assess whether the tablets meet the criteria for the “Uniformity of Weight (Mass)” test.**

The average mass and percentage deviation has been calculated in (c). All the tablets have an average mass of 80 mg or less which means the percentage deviation is 10% and that the Uniformity of Weight test fails if more than two tablets exceed the percentage deviation or when a tablet exceeds double of the percentage deviation which is 20%. We then created a calculated field which helps to identify whether a time point fails or passes the test.



The calculated field checks if any tablet exceeds 20 or if more than two tablets exceed 10% and marks the time point as pass or fail based on these checks. From the results we can see that minutes 270, 315 and 450 did not meet the criteria and failed the test. All the other time points passed the test.

1. **Perform data transformation on appropriate columns. Ensure all transformations are clearly explained, including any formulae or calculated columns, so the data is prepared for precise analysis**

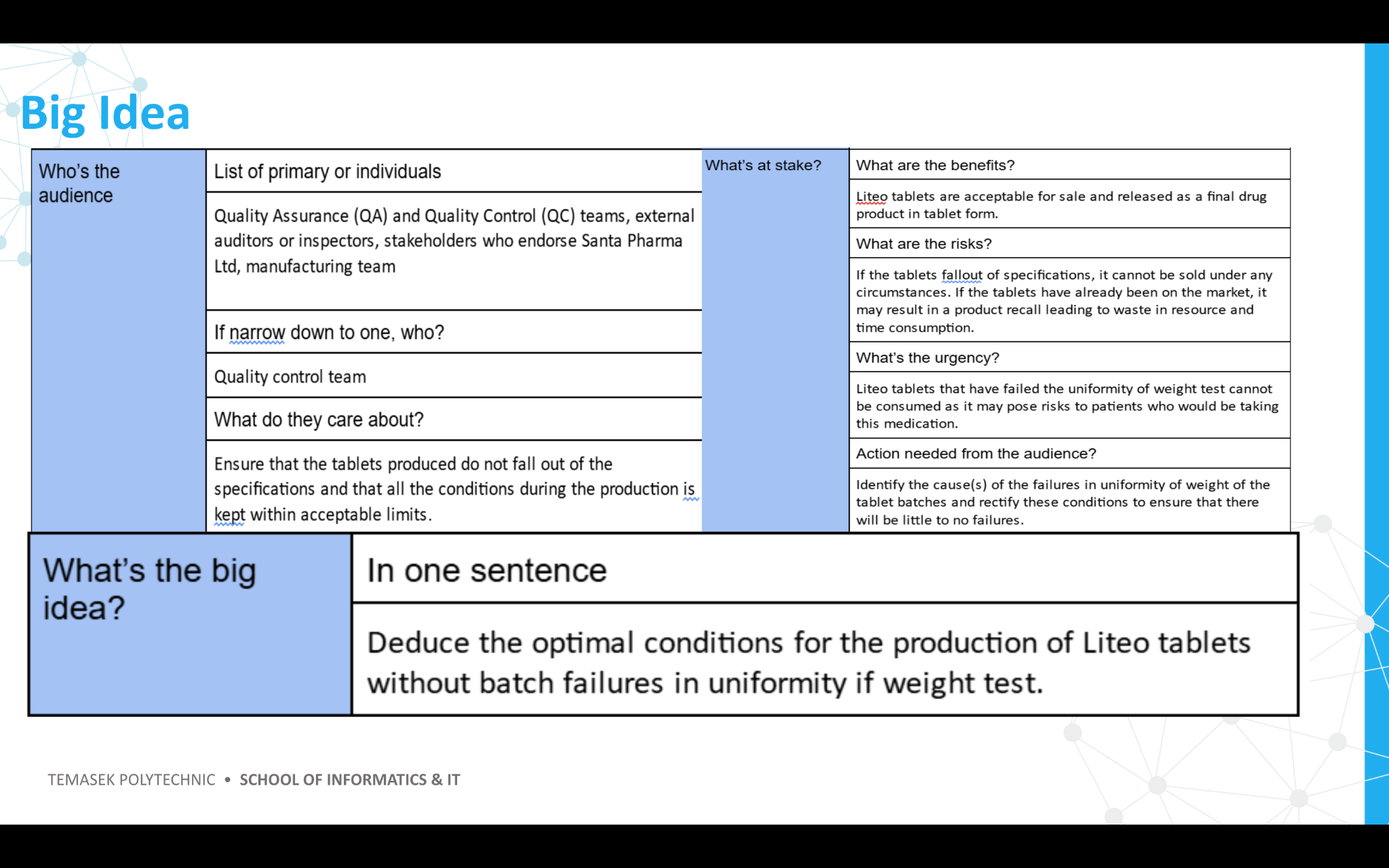
There are no new data transformations that I created, there are only the transformations I created earlier such as the percentage deviation, uniformity test results, average mass and the time value field for the weight, height, relative humidity and temperature.

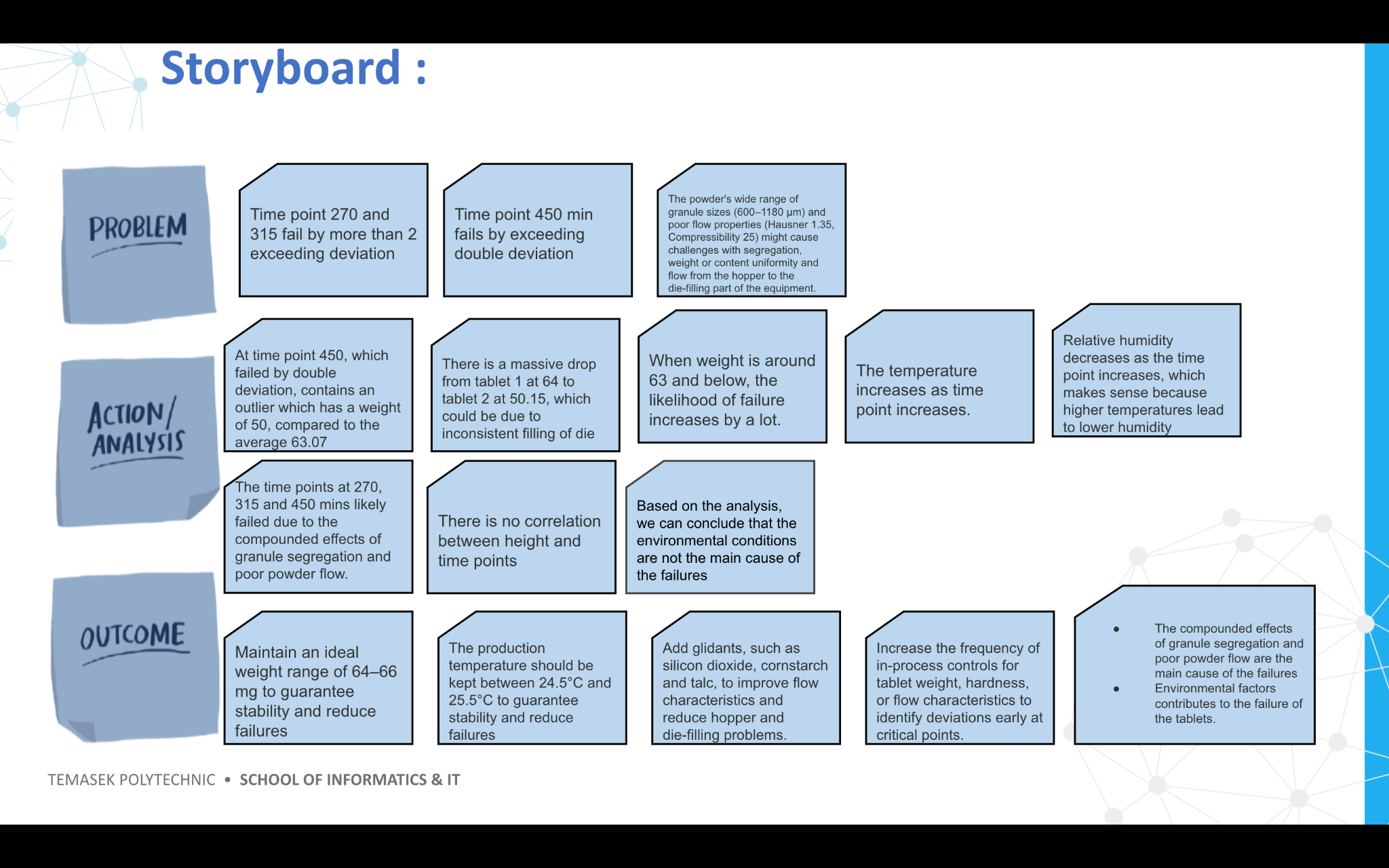
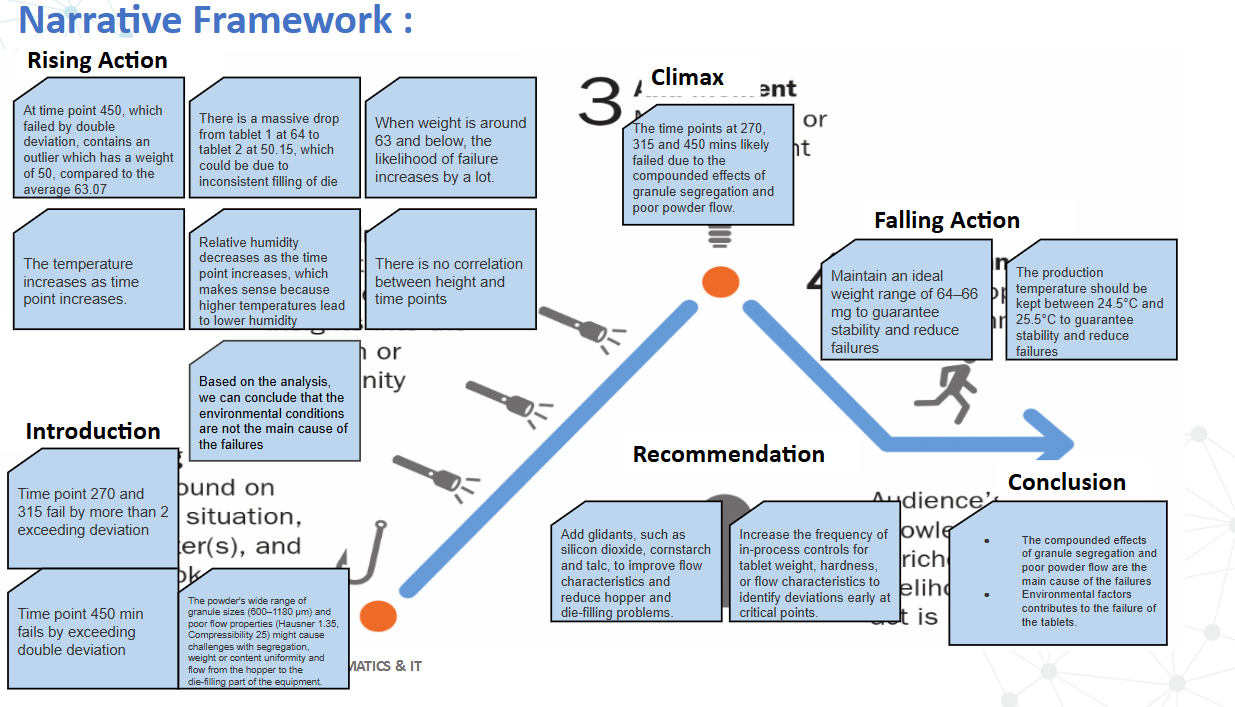
The formula is as followed:

Average Mass = Sum of tablet weights at a time point / Number of tablets

Percentage Deviation = (|average mass - tablet weight| / average mass) x 100%

# **4.** **Data Storytelling**





The introduction establishes the necessity for research by highlighting significant tablet production failures at specific times. By highlighting the importance of the failures, this introduction creates initial tension.

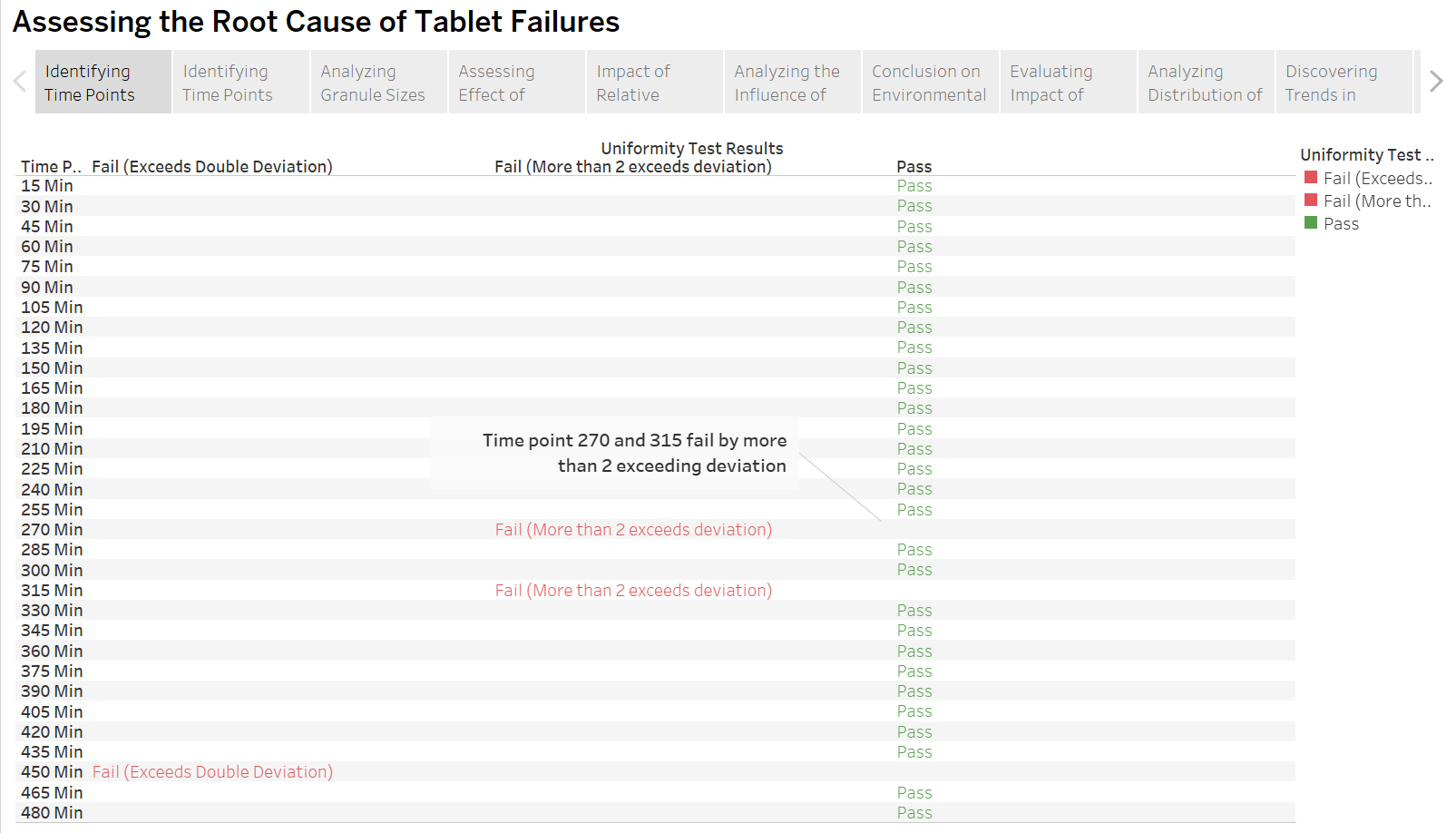
More details are displayed by the rising action, which includes the possible causes of failure such as the environmental factors and the poor powder flow and granule segregation. These details stimulate curiosity and bring us closer to the turning point.

The climax reveals that the primary reasons for failures are granule segregation and poor powder flow, highlighting the necessity of corrective action.

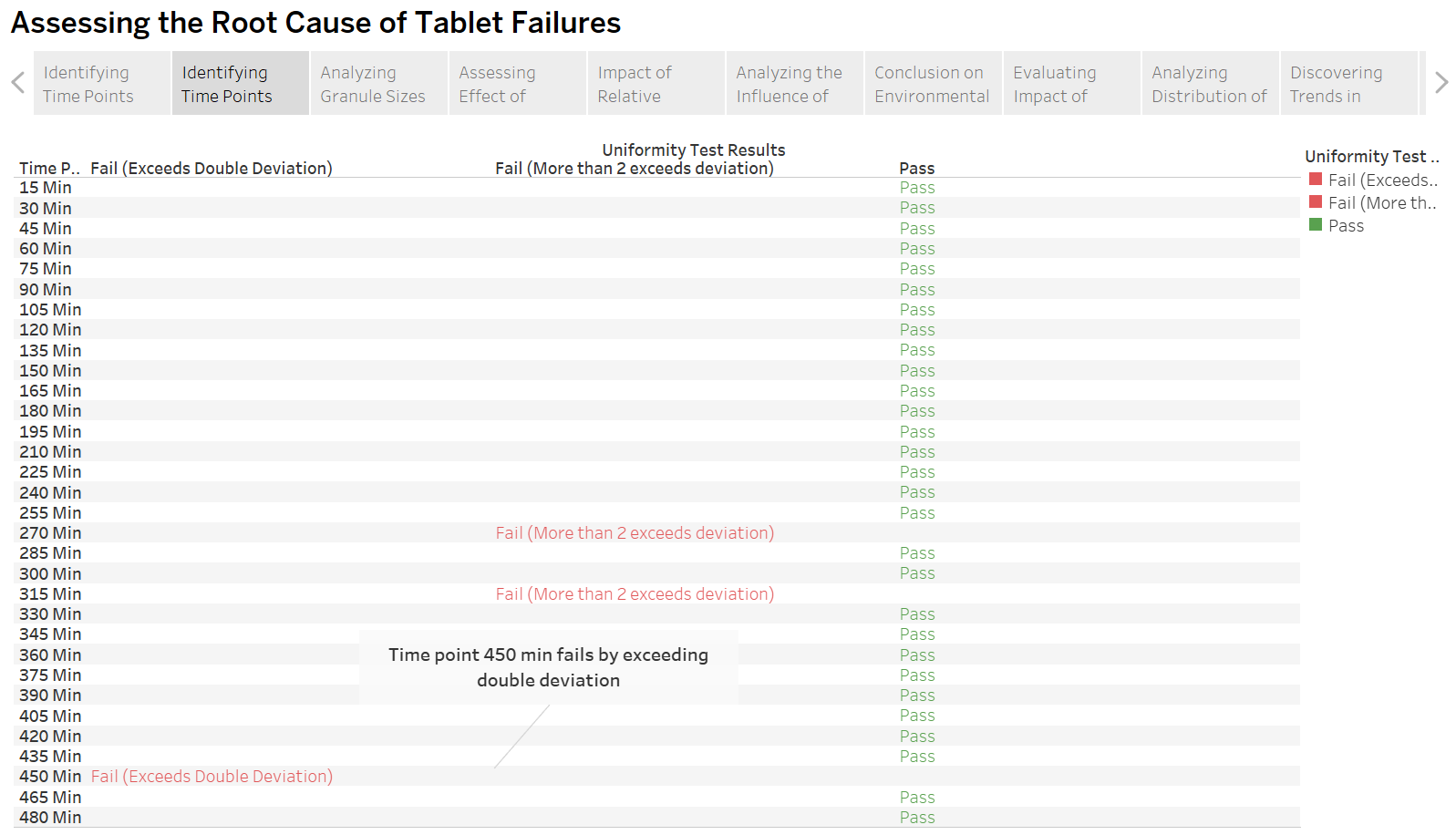
The falling action provides recommendations that could be immediately put into practice to address the problems mentioned in the climax These recommendations include maintaining a stable temperature and weight range to increase manufacturing stability. These feasible solutions are intended to directly address the main issues noted, guaranteeing a prompt and efficient resolution

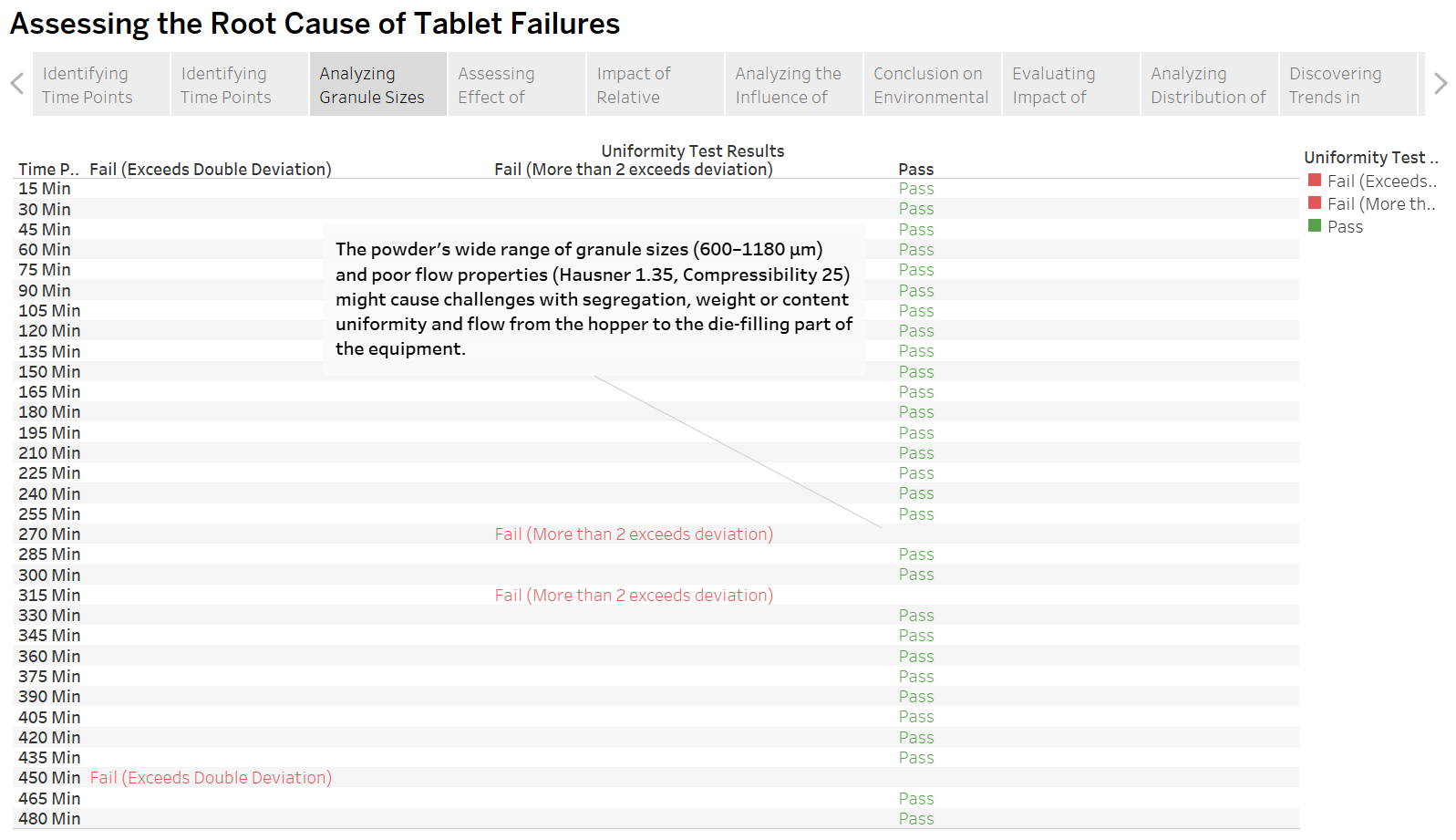
The conclusion reinforces the significance of resolving the challenges raised by bringing the main conclusions and useful recommendations together. These recommendations include enhancing in-process controls to identify deviations early and incorporating glidants to enhance flow for improved production. The conclusion successfully answers the audience's main concerns by outlining the advantages of these steps, responding to the question, "What is in it for me?" and motivating them to implement the suggested fixes.

**Introduction**

The introduction outlines the main production problems faced, such as during time points 270, 315 and 450, where the deviations exceed the allowable range. It makes use of powerful visuals to establish a captivating hook and inspire a sense of urgency to solve these issues identified.

The graph shows the uniformity of weight test and highlights the time points, such as 270 and 315 which failed by more than 2 exceeding deviations.

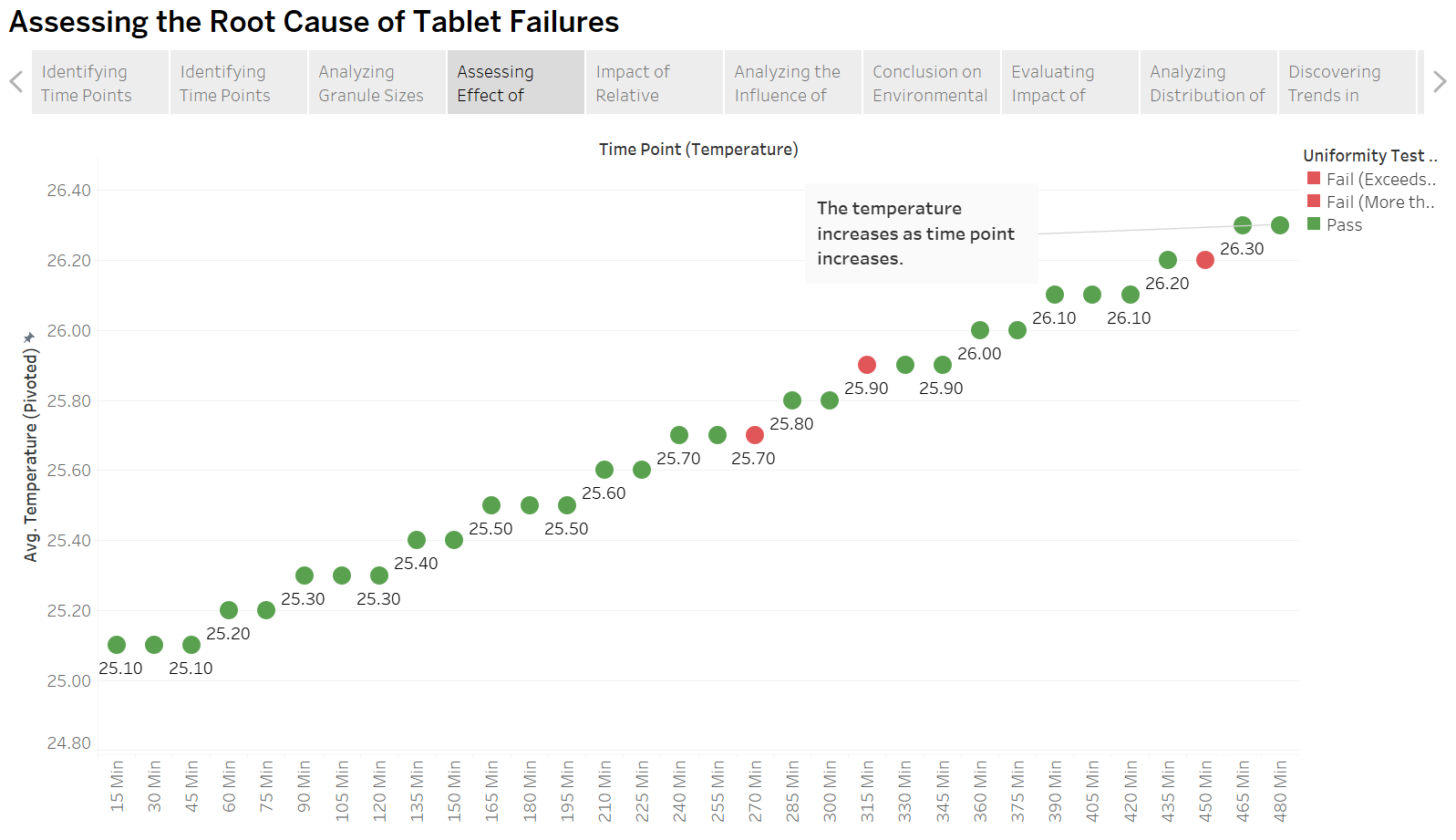
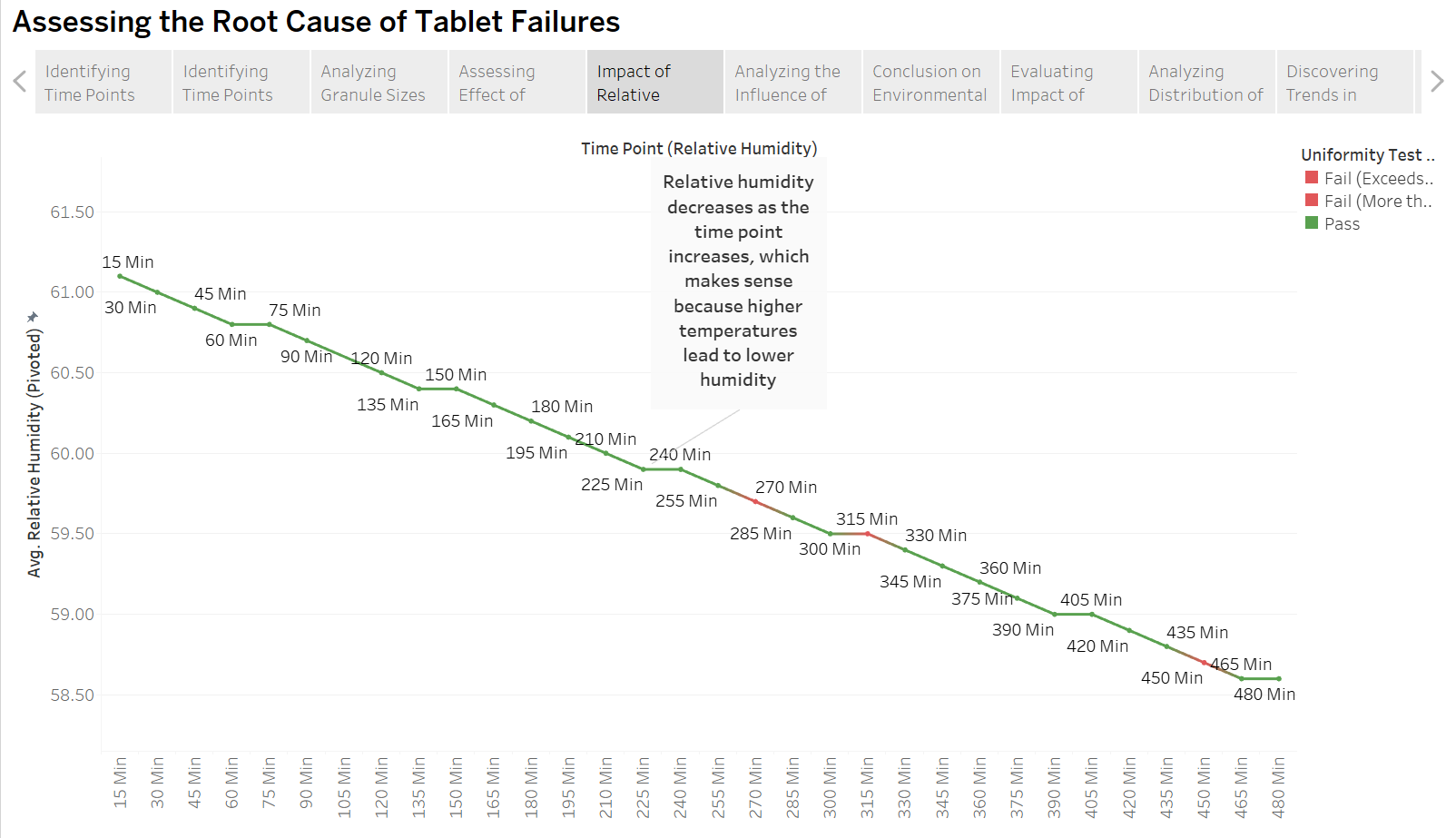
Time points 450 mins failed by exceeding double deviation



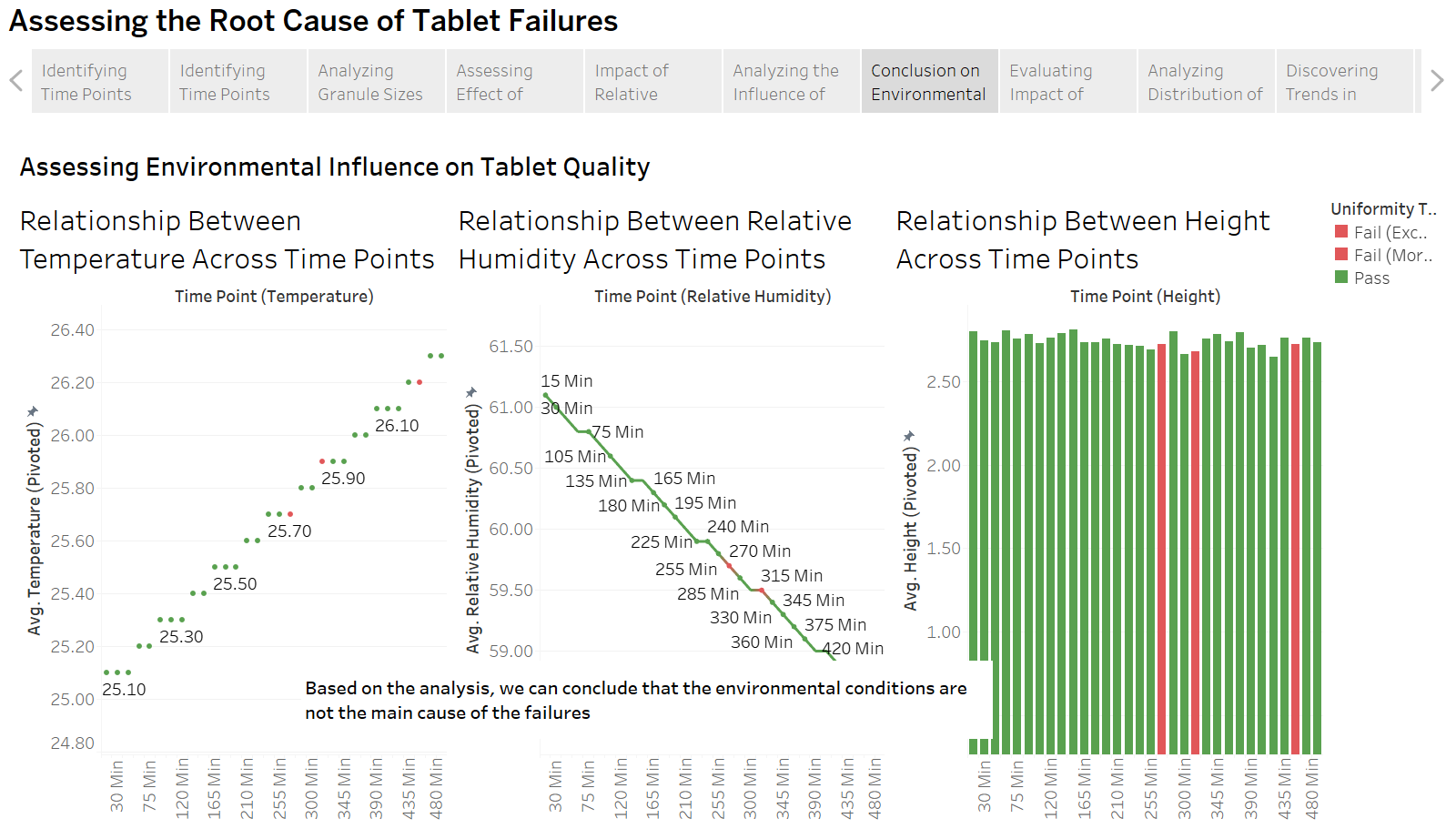
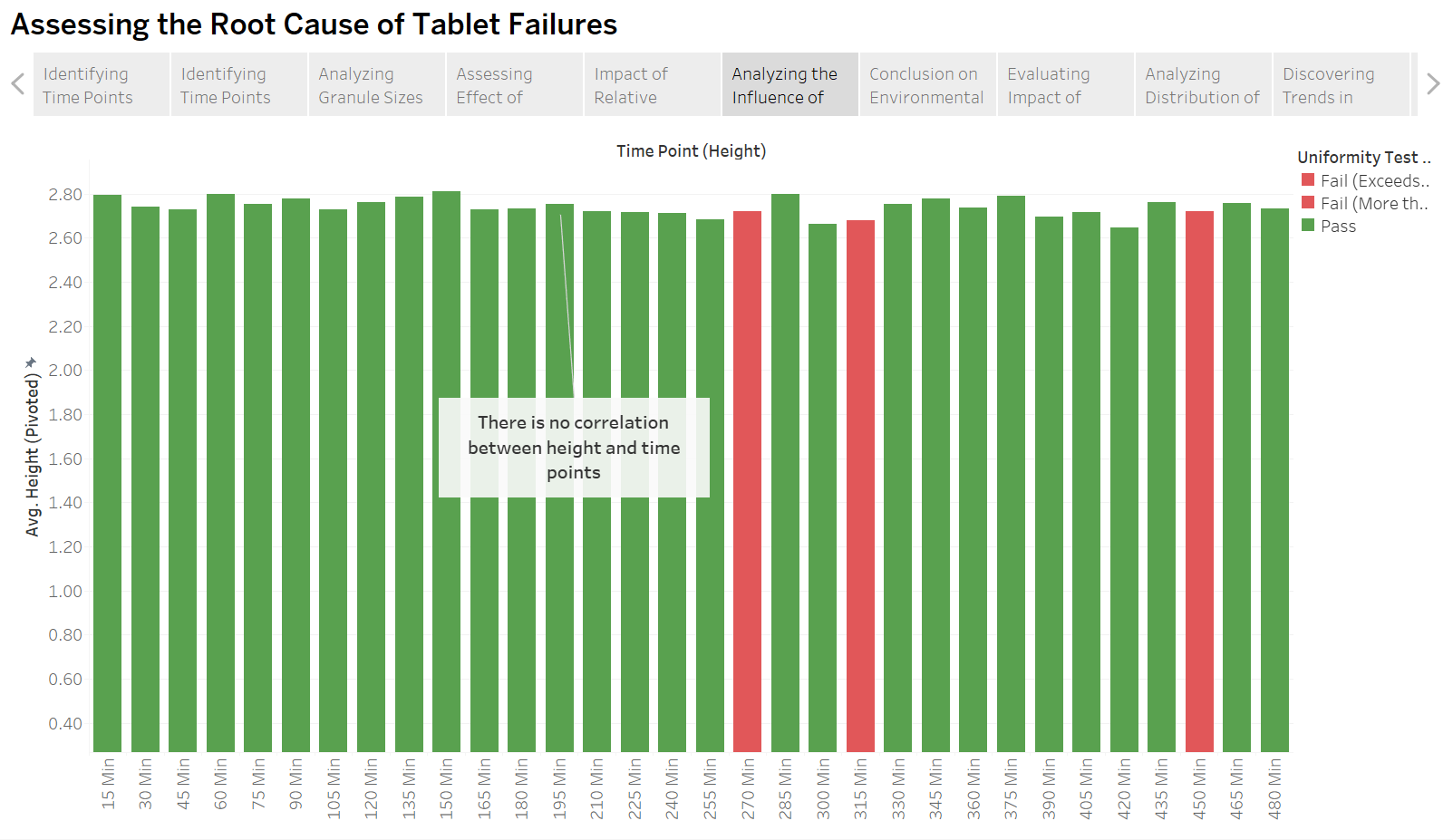
From the background information, we can identify that the tablet properties are poor with a Hausner ratio of 1.35 and compressibility index of 25, with a wide range of granule sizes (600–1180 μm). These poor properties can cause challenges with the segregation, weight or content of uniformity and flow from the hopper to the die-filling part of the equipment.

When combined, these visuals form an engaging narrative framework that successfully introduces the issues in the tablets. The introduction makes sure that the audience is interested and prepared for the more in-depth analysis and solutions that will be provided in the following sections by combining high-level and specific ideas.

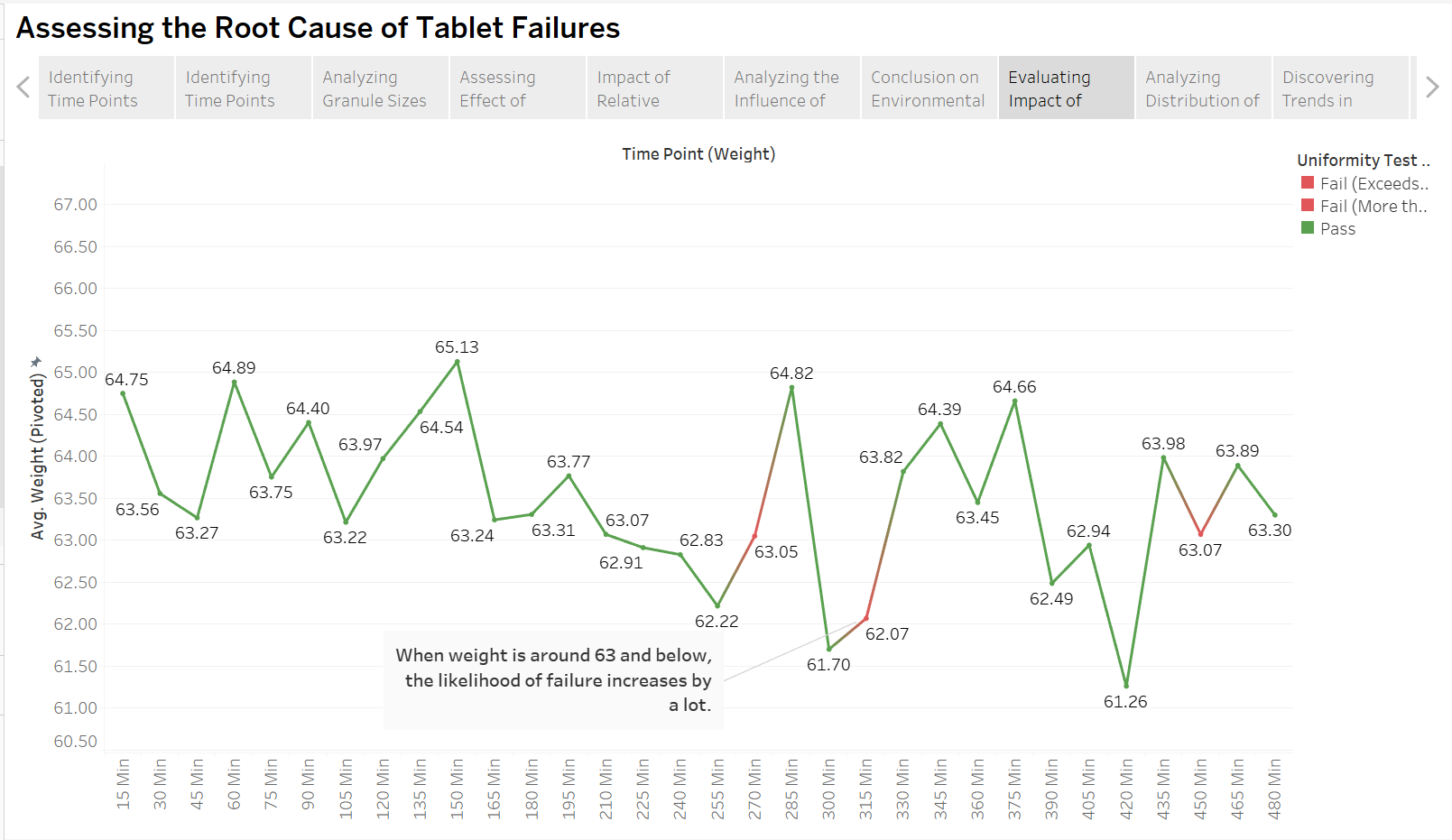
**Rising Action**

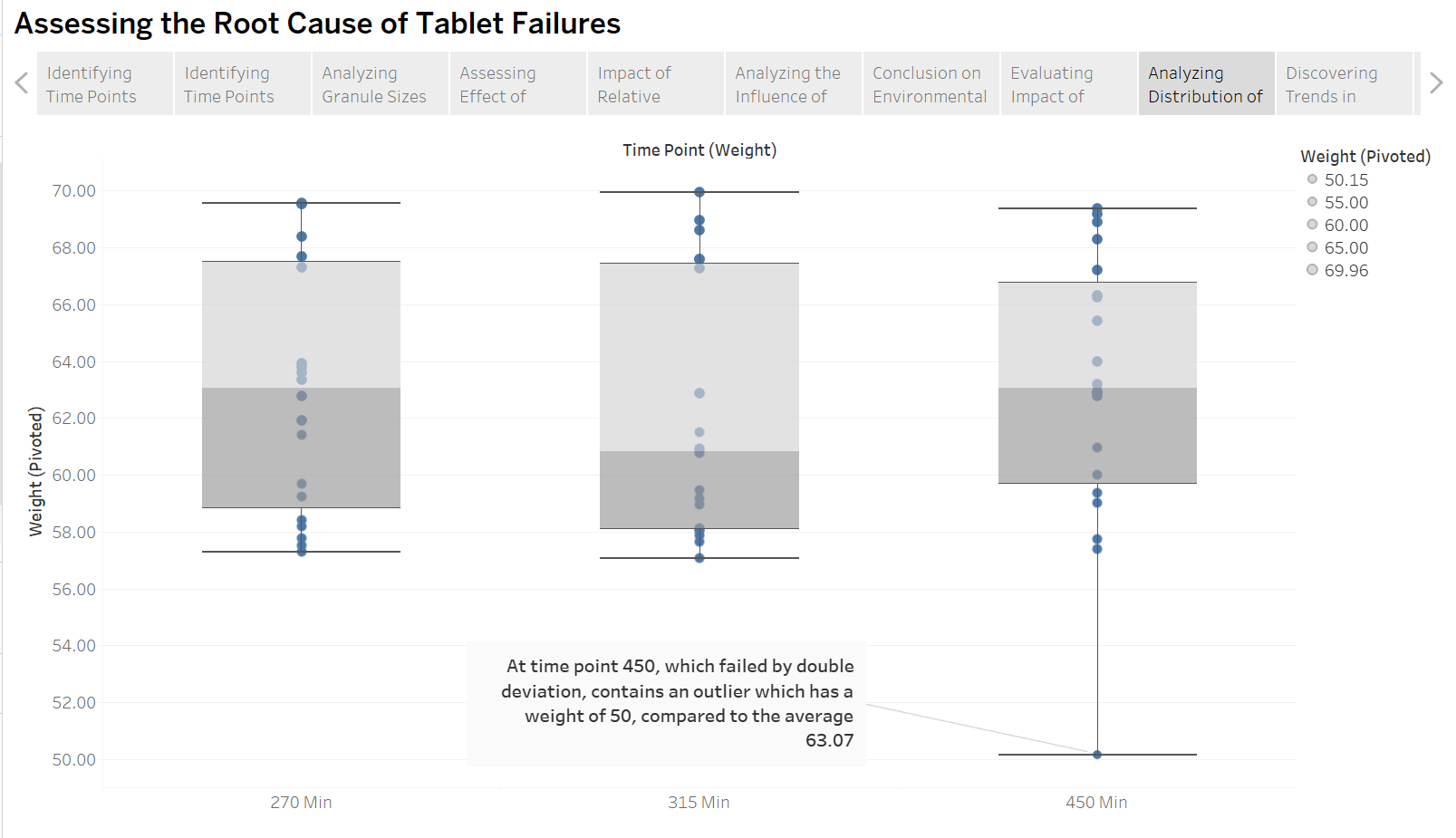
The rising action analyses the possible causes of the failures such as environmental factors and the poor tablet properties. These charts offer intermediate results, laying the groundwork for the main revelation in the climax.

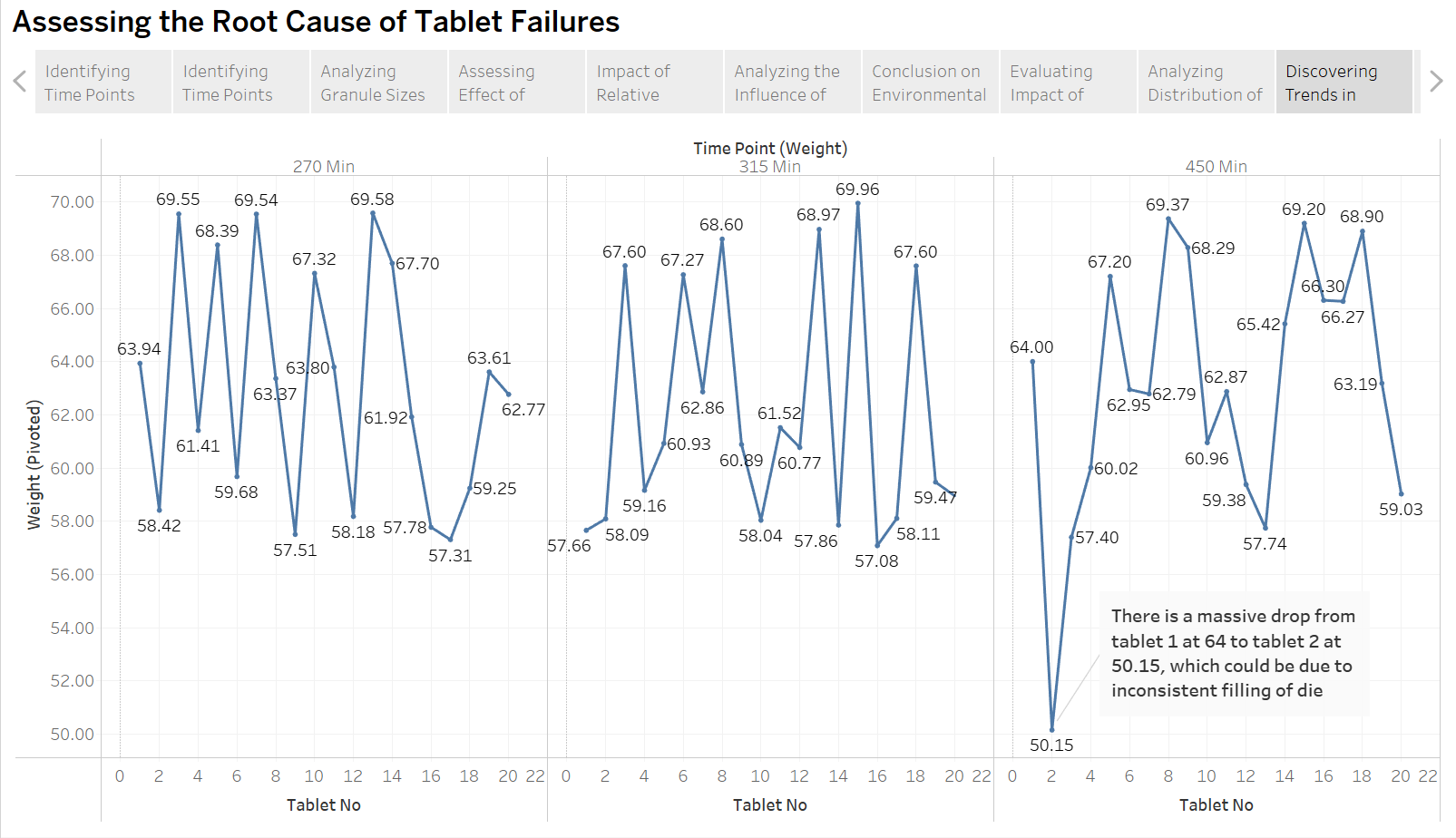
The first chart shows that the temperature increased noticeably during the production process. Over time, the recorded temperature steadily increases, peaking at later points. The second chart shows that humidity decreases with time. This pattern makes sense as rising temperatures lead the air to hold less moisture, which lowers humidity. This relationship is visually supported by the chart, which displays a consistent drop in humidity over time.

There is no correlation between height and time point which means that it does not cause the time point to fail.

This suggests that although environmental factors such as changes in temperature and humidity do contribute to tablet failures, they are not the main cause. Relative humidity falls and temperature rises gradually over time, as the charts demonstrate, but failures happen at times rather than continuously following these patterns. As such, environmental conditions are not the main cause of the failures.



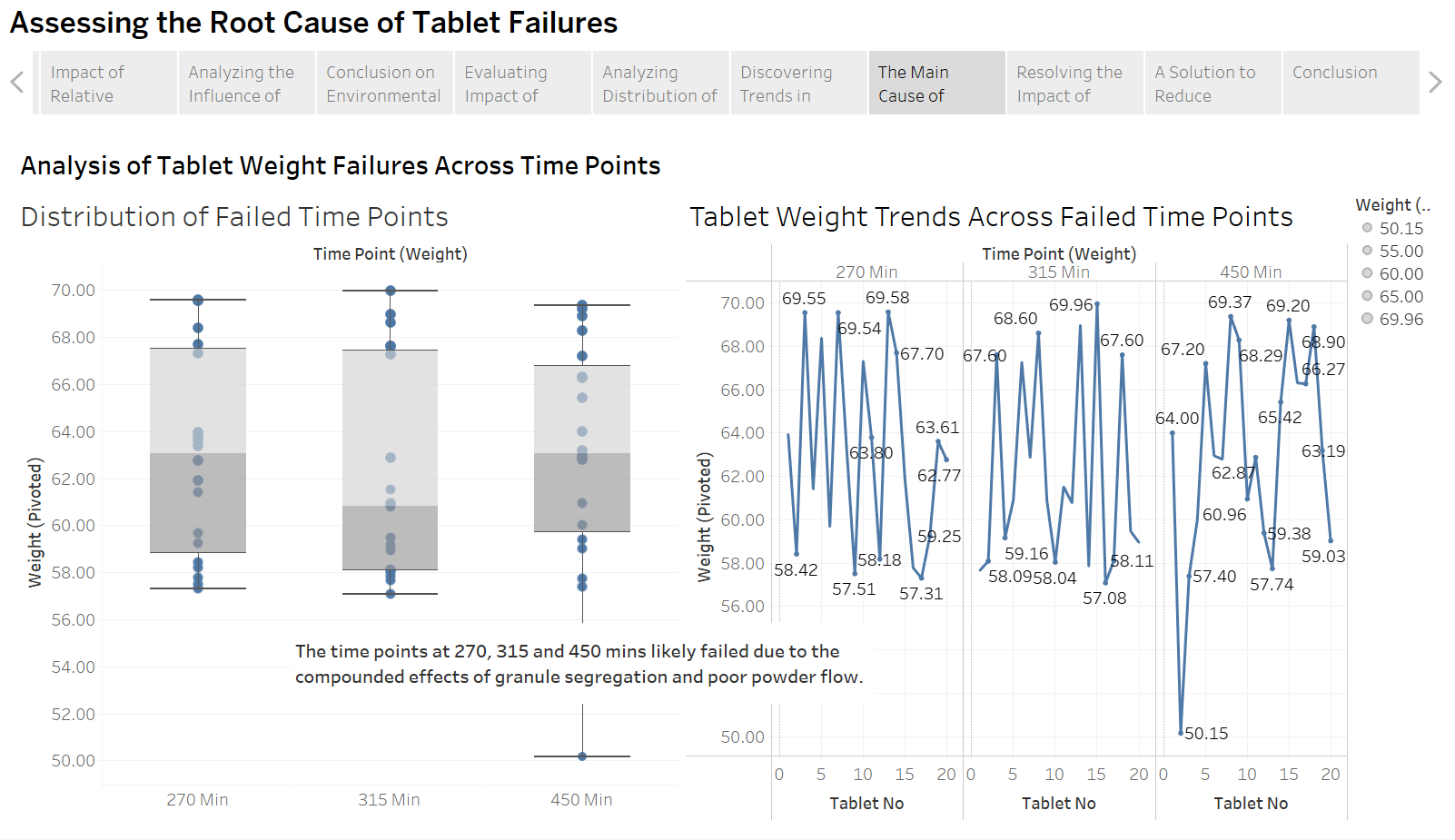
We perform a more thorough analysis to determine the main reason for the failures because environmental factors are not the main cause. The average tablet weight over time is displayed in the graph above. Time points 270, 315, and 450 all failed at 63.05, 62.07, and 63.07 mg, respectively, according to the graph, which shows that all failed time points have weights of 63 mg or less. This suggests that failure rates have a high correlation with decreased tablet weight.

An extreme outlier—a tablet with a much lower weight of 50 mg, when the average weight at this time point is 63.07 mg—was one of the main causes of the tablet production failure, which was more severe (double deviation) at 450 minutes. The outlier is located well below the lower quartile in the box plot, suggesting an uneven weight distribution.

We created a line chart to examine the weight distribution of individual tablets at failure time points after locating the extreme outlier in the box plot. At time point 450, the graph shows a notable decrease in weight from 64 mg at tablet 1 to 50.15 mg at tablet 2. This sharp drop indicates inconsistent die filling, which could have played a part in the failure.

By building tension and preparing the viewer for practical answers at the climax, these rising action visualisations work in unison to show mounting evidence of tablet production issues and support the narrative structure. By offering in-depth, visually appealing insights that link the introduction and the climax, they enhance the story and emphasise how urgent change is.

**Climax**

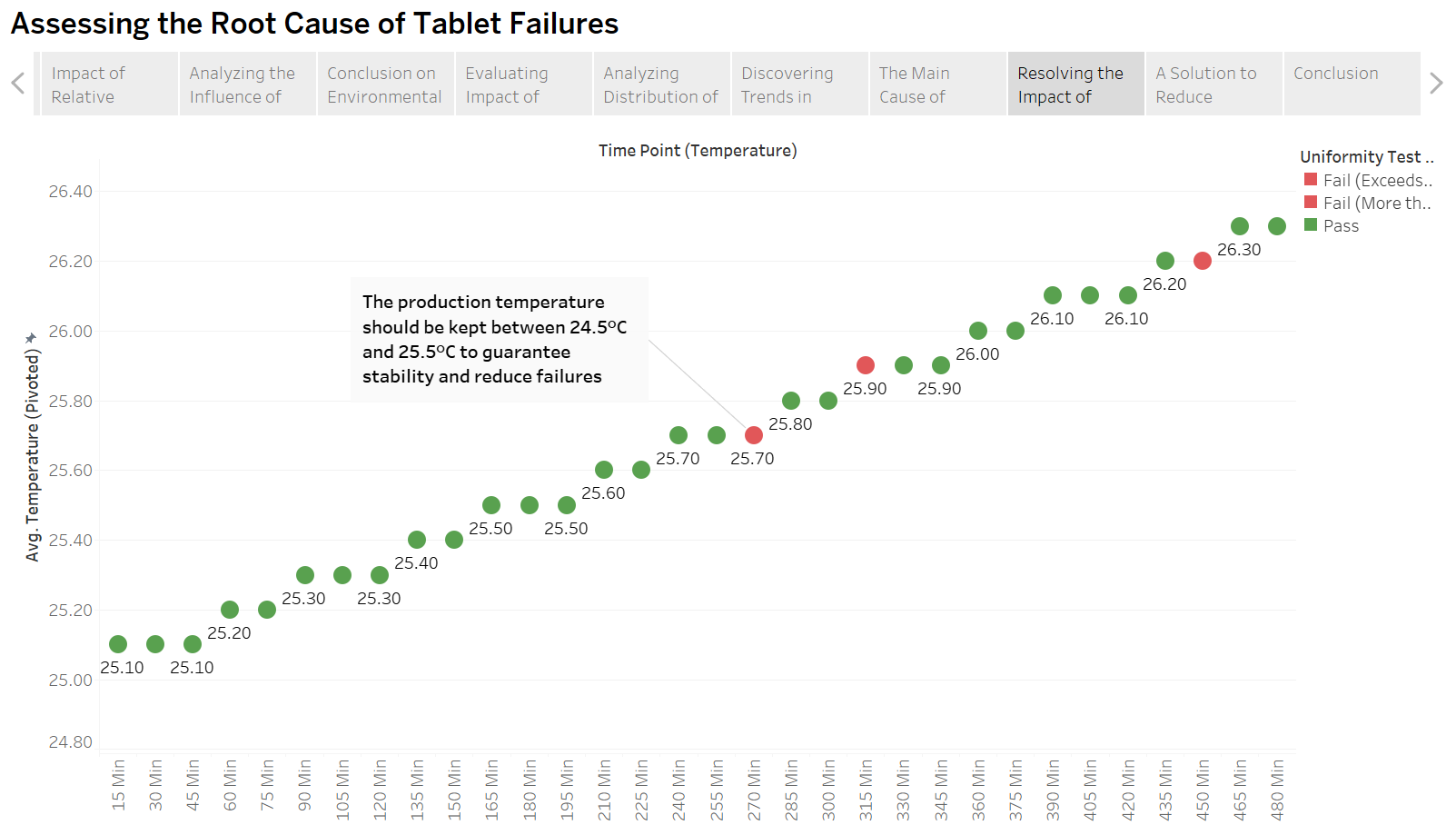


Based on all the analysis we performed and all the information we have gathered, we have concluded that the main cause of failure for time points 270, 315 and 450 mins is due to the compounded effects of granule segregation and poor powder flow.

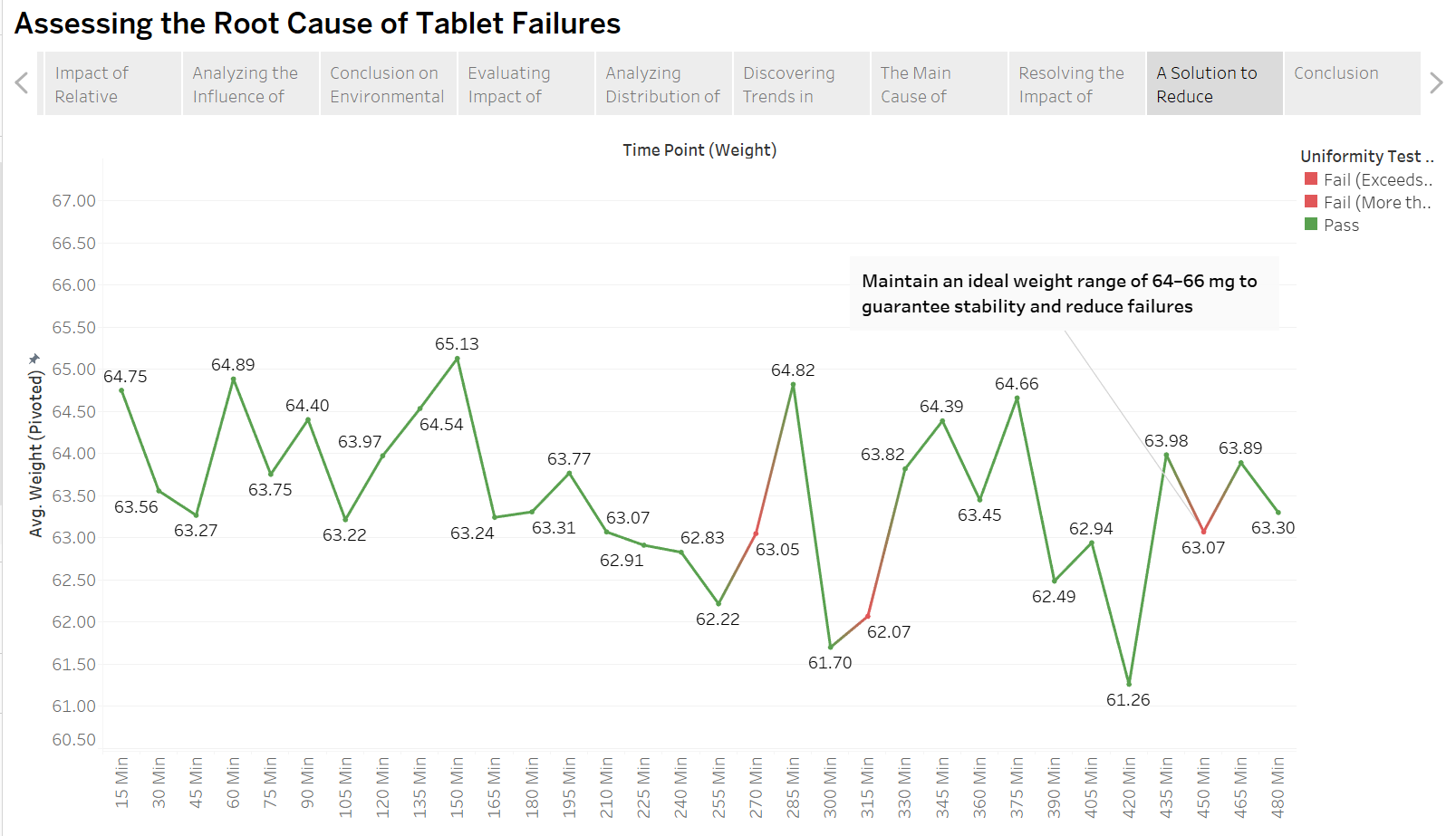
These visuals emphasise the effect of inconsistent die filling on production quality, which cultivates a sense of urgency that fits within the story's structure and acts as a powerful hook. Decision-makers must place a high priority on ensuring the powder flow meets the “excellent” flow character according to USP standards and also standard using the granules to a narrow distribution range to minimise failures in the uniformity of weight. The findings are supported by data-driven insights and visual storytelling, which prompts prompt action to be taken.

**Falling Action**

The falling action visualisations focus on realistic recommendations to address granule segregation, poor powder flow and the environmental factors, ensuring a smooth transition from the climax. With their emphasis on urgency and concise solutions, these charts offer useful insights and reinforce the narrative framework.



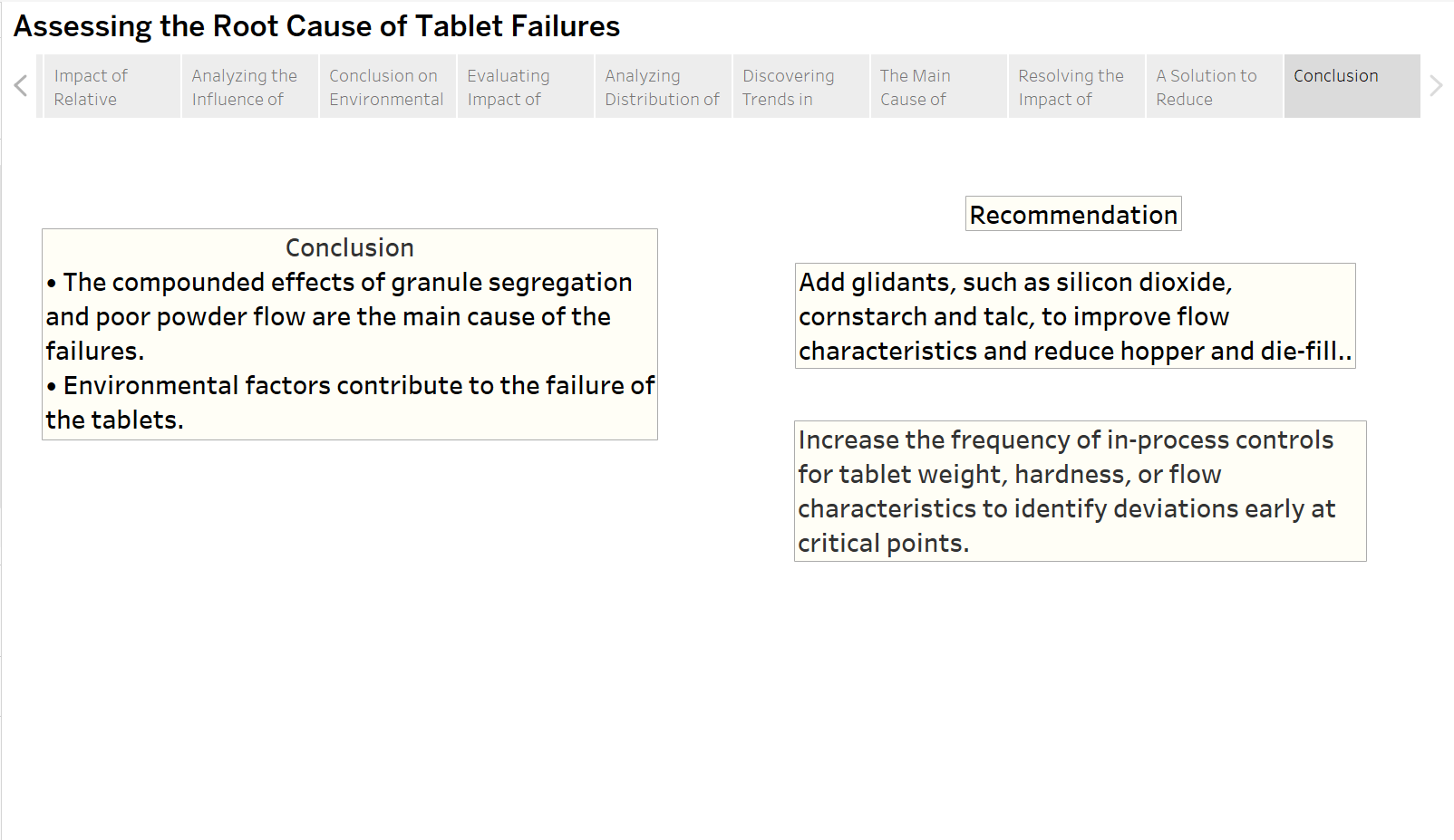
We suggest maintaining the production temperature between 24.5°C and 25.5°C to minimise tablet failures and guarantee process stability. Failures typically occur at higher temperatures, according to our previous findings, which implies that going above this range could lead to inconsistent tablet output. Improving uniformity and reducing flaws can be achieved by maintaining the temperature within the suggested range.



In order to minimise failures and guarantee consistent manufacturing, we suggest maintaining tablet weights between 64 and 66 mg. The necessity for stricter weight management is highlighted by our prior research, which shows that tablets weighing less than 63 mg are more likely to fail. We may increase uniformity, improve product quality, and lessen manufacturing process irregularities by maintaining tablet weight within this range.

These charts give a clear call to action and instil a sense of urgency. By providing practical solutions to tablet production issues, they respond to the problem statement. The suggestions are easy to comprehend and put into practice thanks to the straightforward and understandable visuals. The conclusion follows naturally from this falling action, which successfully connects the insights to solutions.

**Conclusion**

The conclusion summarises the analysis's main findings and offers useful recommendations for resolving the problems that were found. It successfully incorporates the narrative structure by reiterating the key findings, such as the effects of granule segregation, poor powder flow, and environmental factors.

By immediately connecting the conclusion to the suggestions, the visual structure reinforces the message and guarantees a clear course of action. The recommendations centre on enhancing in-process controls to identify deviations early and adding glidants (such as silicon dioxide, cornflour, and talc) to improve powder flow to improve tablet production quality.

This conclusion highlights how urgent it is to address these problems and increases the confidence in the suggested solutions. By acting as a clear call to action, it guarantees that the required actions are performed in order to reduce faults, improve manufacturing efficiency and also ensure safety in the tablets produced that might be set out for sale.

**Appendix: Reflective Journal**

Answer the following questions:

1. Rate the usefulness of Gen AI tools used by your group (choose one) : 1 / 2 / 3 / 4 / 5

2. Tell us how your group uses Gen AI tools in your IDL assignment. Please include some screenshorts/ prompts used.

3. Did you encounter any challenges when using Gen AI tools in this assignment?

4. Can you provide some suggestions for future implementations?

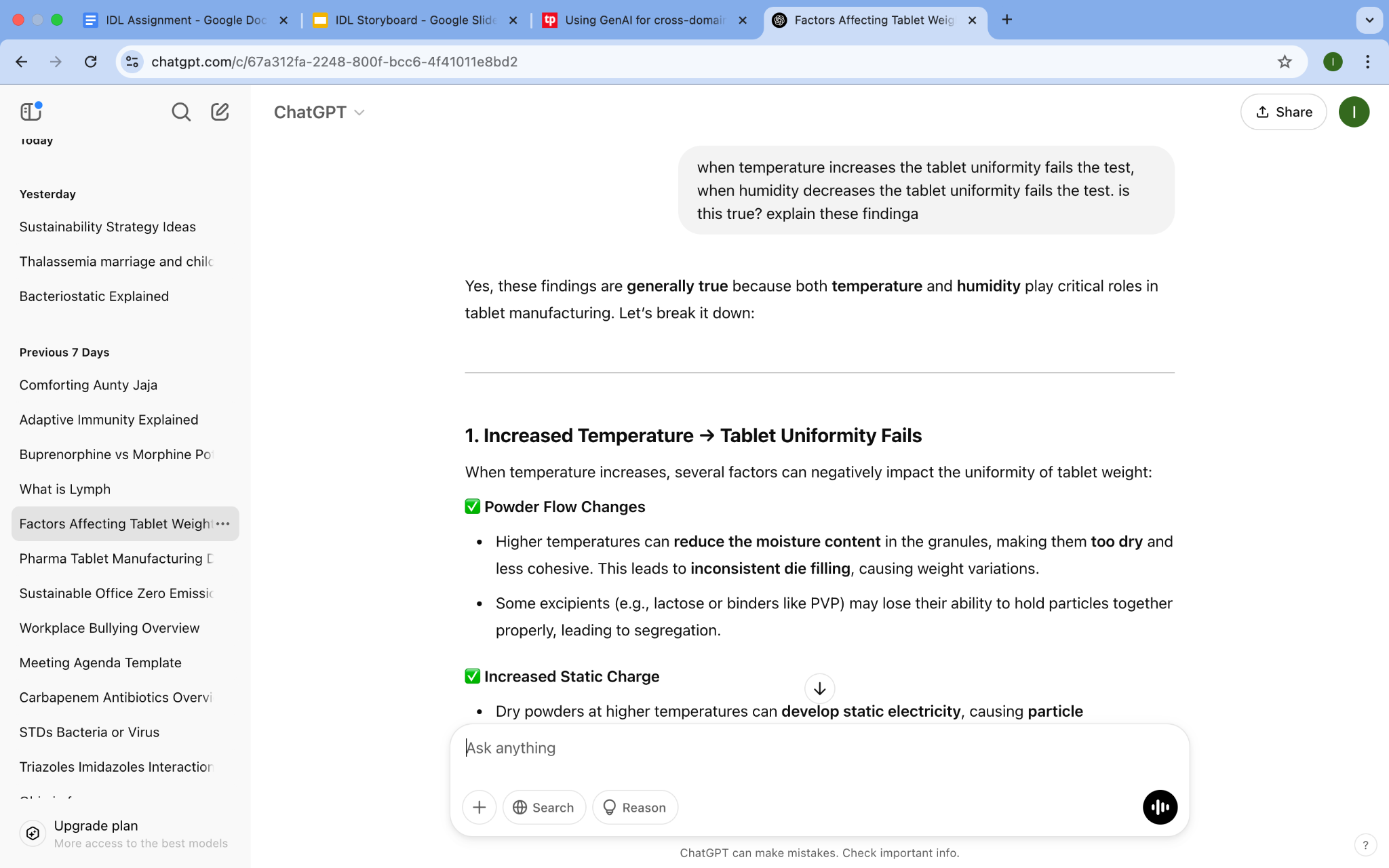
1. 3
2. We utilized Generative AI tools to assist in analysing certain aspects of our results that were unclear, specifically the results from our graphs. Gen AI provided valuable insights about the trends and the reasons as to why the results were as so, helping us better understand the factors influencing our data. It also was able to provide us with a scientific explanation, where we were able to link that information to the principles of pharmaceutical manufacturing. However, when we couldn’t trend and ran out of ideas for us to give a reasoning to our charts, we managed to input the results and they helped to generate other possible reasons which to me made sense and were pretty accurate to the scientific fact.

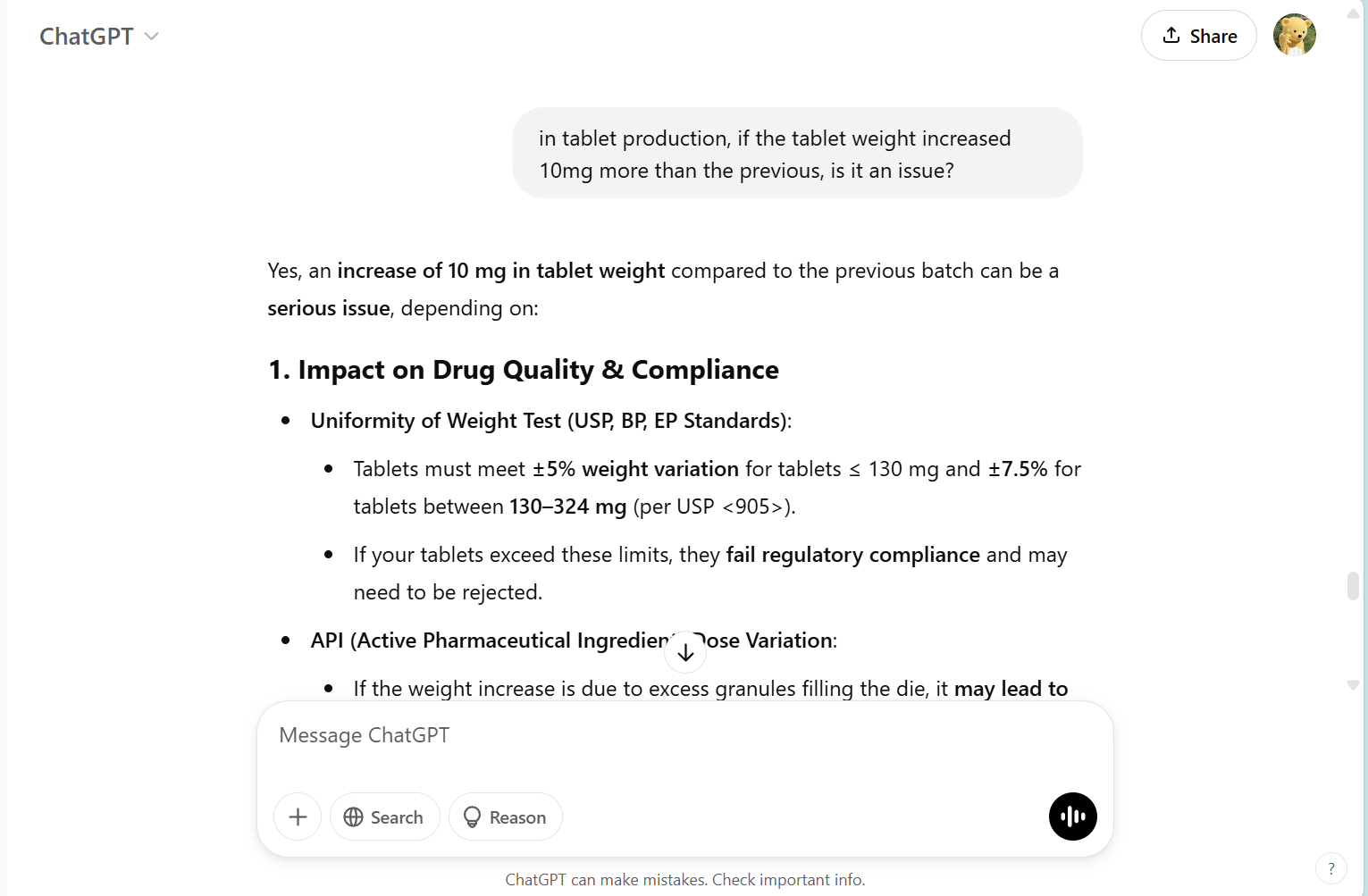
Beyond helping us analyse our results, Gen AI also helped to break down the effects in a comprehensive way. By clarifying how temperature and humidity affects tablet uniformity, the explanation enhanced our understanding and provided a well-supported conclusion to our findings.

If the terms were too difficult for other group mates to understand, we could prompt them to generate a more laymen term way of explanation hence it was easier to understand the concept after the prompt had been modified in this way.

Additionally, the Gen AI was able to give us an analysis with reference to the different pharmacopoeias without us needing to individually go through each one which saved a lot of our time with the project.

Overall, Generative AI has offered us clearer insights, scientific reasoning, and a structured breakdown that strengthened our understanding and discussions.





1. One challenge we encountered was applying the findings and the information presented by Gen AI into the specific context of our results. Although the information from Gen AI was concise and detailed, we had to strategically evaluate its responses, ensuring its validity and accuracy. Another challenge that we encountered was when we needed specific and precise answers, Gen AI tends to provide more general and lengthy responses. However, we were able to overcome this issue by refining our prompts, specifically letting the AI know what we needed. I would also like to add that the responses can be quite repetitive even after modifying the prompts which can be a little challenging if we are trying to find more answers but the AI keeps repeating the things we have already noted down.
2. Some suggestions for future implementations:

* Using Gen AI in combination with assignments, where we have to use Gen AI and cross-check it’s accuracy and validity with definitive literature.

|  |
| --- |
| SUBMISSION INSTRUCTIONS    · Save this file as “IDL\_Group[Group Number]\_[Group Leader’s Name]\_[Group Leader’s ID]”, e.g. “***IDL\_Group99\_SUPER HERO\_2299999A***”.  · Submit your report to DAST LMS site > Assessment > IDL Assignment Submission Link. |