Python-hackathon presentation

On this slide, we outline the key challenges in working with the HUPA-UCM diabetes dataset and the strategies we can adopt to address them.  
**Challenges:**

1. **Small sample size:** With only 25 patients, it’s difficult to generalize findings to a wider population.
2. **Short monitoring period:** The data covers 14 days or more, which may not capture long-term patterns in glucose or behavior.
3. **Missing or noisy data:** Continuous glucose monitoring and wearable devices often have gaps or inconsistent readings.
4. **High variability in patient behavior:** Differences in lifestyle and treatment responses make modeling more complex.
5. **Complex relationships:** Glucose levels are influenced by multiple interacting factors such as insulin, meals, and activity.

**Strategies:**

1. **Data augmentation or combination:** We can increase robustness by augmenting the dataset or combining it with other Type 1 Diabetes datasets.
2. **Robust preprocessing:** Handling missing values, outliers, and noise to ensure clean data for analysis.
3. **Personalized modeling:** Building patient-level models before attempting group-level predictions improves accuracy.
4. **Feature engineering:** Incorporating features like variability and time-of-day effects helps capture complex patterns.

Overall, by carefully addressing these challenges, we can improve the reliability and predictive power of our models**Slide 4**On this slide, we focus on **prescriptive insights and recommendations** derived from our analysis.  
First, **maintaining a consistent sleep schedule** helps stabilize glucose levels, which is crucial for overall metabolic health.  
Next, **balancing carbohydrate intake with bolus insulin** can reduce glucose spikes after meals, supporting better glycemic control.  
We also recommend **encouraging daily activity**, such as tracking steps, which improves insulin sensitivity and overall metabolic efficiency.  
Finally, looking ahead, there is significant **future scope** in developing **personalized digital health monitoring and decision-support tools**. These can provide individualized recommendations to optimize glucose management and lifestyle choices.  
Overall, these actionable insights can help individuals with diabetes make informed decisions to maintain better glucose control and improve their overall health**Slide 3**  
let's walk through what we found after diving deep into the data.  
**First up, we had to make sure we were working with a clean dataset.** We scrubbed the numbers by removing any duplicate entries, filling in those pesky missing values, and—importantly—correcting some invalid insulin dose recordings. This means the insights we're about to go through are built on a solid, trustworthy foundation.  
**Now, onto what the data is actually telling us.** A few key themes emerged from our descriptive analysis:

* We saw that **glucose variability is notably high** across our patient group. This isn't just a minor fluctuation; it's a significant pattern that we need to pay attention to.
* Digging into why, we discovered a **strong, positive correlation between carbohydrate intake and insulin dose**. The correlation coefficient sits at about 0.46, which is a substantial relationship. Simply put, the more carbs a patient consumes, the higher their insulin dose tends to be.
* And it's not just about food. We also identified a clear link between **sleep irregularities and glucose fluctuations**. When sleep patterns are off, glucose levels tend to be more volatile, which is a crucial factor for patient management.

**Finally, we used this cleaned data to look forward.** Our predictive analysis involved building regression and correlation models. These models are powerful because they allow us to **forecast future glucose trends** for patients. Even more critically, they've helped us **identify the key risk factors that make a patient more susceptible to hypo- or hyperglycemic events.**  
So, in summary: we've cleaned the data, uncovered the core drivers of glucose variability, and built tools to not just understand, but anticipate patient risk.

Dashboard:  
A dashboard is a visual interface that consolidates key information and metrics in one place, enabling quick insights and decision-making. This dashboard provides a visual summary of key patient data, combining multiple charts in one view for easy analysis. It allows us to quickly understand age distribution, sleep disturbances by age group, and race composition, helping identify patterns and insights at a glance.