**MILESTONE 1 (15 minutos)**

1. **Team Roles (who):** (podríamos pedirle a Jose los resultados de nuestra encuesta y ya definimos a partir de ahí: 1 persona)

**::::::::::::::::::::::(Inglés)**

First of all, with the aim of having a harmonious team where everything fits and works correctly, we decided to assign certain roles. However, these roles did not determine the area where each member worked. They have determined how we have managed the tasks.

For the team roles, we mainly took into account the test results. For that reason, the leader of the group is Manuel, the doer is Yassmina, the thinker is Evgeny and the carers are Jezabel and Carla.

And how has this impacted our sessions? Well, the leader took the initiative and organized overall, the doer took notes and primarily drafted, the carers maintained harmony by assisting in all areas and planning, and the thinker provided ideas to the team and helped with weekly tasks.

1. **Context:** (añadiría este punto explicando de qué va nuestro tema y demostrando que controlamos el ámbito: 1 persona)

WHAT, Why, How (este punto sería uno de los what)

EXPLICAR LAS COSAS COMO UN MEDICO

Meterle miedo a la gente

Porcentaje de la cantidad de gente que lo sufre

Marcar las dos partes (Dentoalveonar y peri-implantitis)

Fotos asquerosas

En concreto, tras evaluar los datos de 474 implantes de 275 pacientes, el estudio desvela que la prevalencia de enfermedades periimplantarias en la población española adulta es del 51% (un 27% con mucositis y un **24% con periimplantitis**).

**::::::::::::::::::::::(Español)**

Nuestro proyecto se enfoca en **mejorar la eficiencia y calidad de la atención en Cirugía Dental.**

* Investigamos los **factores que afectan el tiempo operatorio** para prever la duración de consultas individuales, **optimizando** la **gestión clínica** y la programación de pacientes. Esto nos permite **asignar cirujanos más experimentados** a pacientes que puedan presentar complicaciones identificadas mediante nuestros análisis.
* Además, **identificamos perfiles de riesgo para complicaciones postoperatorias**, lo que permite un monitoreo riguroso y una mejor gestión de costos.

Nuestros análisis tienen como objetivo **descubrir ideas valiosas para los dentistas**, como correlaciones entre **material quirúrgico y características del paciente** con el éxito del implante.

Nuestra misión es **elevar el nivel de atención, reducir gastos y mejorar la planificación del servicio** para proporcionar una atención completa y bien informada a nuestros pacientes.

**::::::::::::::::::::::(Ingles)**

Our project focuses on **improving the efficiency and quality of care in Dental Surgery**.

* We investigate the **factors affecting operative time** to predict the duration of individual consultations, **optimizing clinic management** and **patient scheduling**. This allows us to **allocate more experienced surgeons** to patients who may present complications identified through our analyses.
* Additionally, we **identify risk profiles for postoperative complications**, enabling rigorous monitoring and better cost management.

Our analyses aim to **uncover valuable insights for dentists**, such as correlations between **surgical materials and patient characteristics** with implant success.

Our mission is to **enhance the level of care, streamline expenses, and refine service planning** to offer comprehensive and well-informed care to our patients.

1. **Goals & Value (what and/or what for):**

**::::::::::::::::::::::(Inglés)**

In our project, we have two main objectives. The first is to study which factors are associated with an increase in operative time in Dentoalveolar Surgery patients. With this, we will be able to predict the consultation time for each client, achieving better management of the clinic.

The second main objective is to identify the profile of people who may have postoperative problems, taking more exhaustive control and maniging the costs of the operations since these types of operations are private.

Moreover, through our analyses, we can find useful and interesting relationships for dentists, such as associations between materials or patient characteristics and implant fixation.

These pórposes have a general objective, which is to improve the quality of care and surgery at the dental clinic, as our studies have repercussions on the costs associated with treatment and on service planning, such as optimizing consultations and preparing resources. Additionally, the patient will be attended to with the necessary time for their own issues and will be informed of the risks involved in their operation.

**::::::::::::::::::::::(Español)**

En nuestro proyecto tenemos dos objetivos principales. El primero es estudiar qué factores se asocian con un aumento del tiempo operatorio en pacientes de Cirugía Dentoalveolar. Con esto podremos predecir el tiempo de consulta de cada cliente, consiguiendo una mejor gestión de la clínica.

El segundo gran objetivo es identificar el perfil de las personas que pueden tener problemas postoperatorios, llevando un control más exhaustivo y gestionando los costes de las operaciones ya que este tipo de operaciones son privadas.

Además, a través de nuestros análisis, podemos encontrar relaciones útiles e interesantes para los dentistas, como asociaciones entre materiales o características del paciente y la fijación de implantes.

Estos propósitos tienen un objetivo general que es mejorar la calidad asistencial y quirúrgica en la clínica dental, ya que nuestros estudios repercuten en los costes asociados al tratamiento y en la planificación de los servicios, como la optimización de las consultas y la preparación de recursos. Además, el paciente será atendido con el tiempo necesario para sus propios problemas y se le informará de los riesgos que implica su operación.

1. **Innovation (why):** (lo entiendo cómo el qué vamos a aportar o qué contiene nuestro trabajo de interesante y algo que hagamos que pueda ser innovador, quizás lo de demostrar la correlación entre colesterol alto o anemia y el fracaso en los implantes dentales: 1 o 2 personas)

**::::::::::::::::::::::(Español)**

Respecto a las razones por las que establecimos esos objetivos, por un lado, lograr el objetivo de tener un modelo que pueda predecir cuánto tiempo podría durar una intervención sería una innovación significativa. Actualmente, a menudo experimentamos retrasos de al menos media hora debido a una organización deficiente del horario o complicaciones durante las intervenciones. Estas complicaciones podrían estar directamente relacionadas con ciertas características de nuestros pacientes, como la posición de sus dientes. Aún estamos trabajando para demostrar este punto, pero todo indica una confirmación de esta hipótesis.

Por otro lado, identificar intervenciones riesgosas basadas en las características físicas y/o patológicas de los pacientes podría ser extremadamente útil. Por ejemplo, podríamos demostrar que problemas como el colesterol alto o la anemia pueden ser factores determinantes. De hecho, hablando con el dentista responsable, el impacto del colesterol alto, por ejemplo, aún es incierto. Por lo tanto, probar tales correlaciones sería tanto innovador como práctico.

(resumen)

Respecto a las razones por las que establecimos esos objetivos, por un lado, tenemos lograr el objetivo de tener un modelo que pueda predecir cuánto tiempo podría durar una intervención sería una innovación significativa, con lo que podríamos reducir el retraso de la clínica .

Por otro lado, identificar intervenciones periimplantarias con riesgos podría ser extremadamente útil, para poder evitar complicaciones. Además, sabemos que por el pequeño numero de datos de este tipo de operación, podemos aportar información que aun no se ha estudiado util e innovadora.

regarding the reasons why we establisd (zous) those objectives, on the one hand, we aim to achieve the goal of having a model that can predict how long an intervention might last, which would be a significant innovation, allowing us to reduce clinic delays.

On the other hand, identifying peri-implant interventions with risks could be extremely useful in order to avoid complications. Additionally, we know that due to the small amount of data on this type of operation, we can provide information that is still unexplord but useful and innovative.

**::::::::::::::::::::::(Inglés)** PROPUESTAS PARA RESUMIR O RECORTAR

Regarding the reasons why we set those objectives, on the one hand, achieving the objective of having a model that could predict how much time an intervention could last would be a significant innovation. Currently, we often experience delays of at least half an hour due to poor schedule organization or complications during interventions. These complications could be directly related to certain characteristics of our patients and the position of their teeth. We are still working to prove this point, but everything indicates a confirmation of this hypothesis.

On the other hand, identifying risky interventions based on the physical and/or pathological characteristics of the patients could be extremely useful. For example, we could demonstrate that issues like high cholesterol or anemia can be determining factors. In fact, speaking with the responsible dentist, the impact of high cholesterol, for example, is still uncertain. Therefore, proving such correlations would be both innovative and practical.

1. **Data (with):** (diría el origen o la fuente de datos, breve descripción de la base de datos, los años y demás y qué transformaciones hemos hecho: 2 personas)

**::::::::::::::::::::::(Español)**

En cuanto a los datos, para llevar a cabo nuestro proyecto, el responsable de una clínica dental en Barcelona, Rui Figueiredo, nos proporcionó algunos datos. Rui es un dentista de renombre en Barcelona y continúa ayudándonos enormemente. Los datos consisten en tres conjuntos de datos: uno de 2016, que tiene 1159 observaciones; otro de 2017, con 719 observaciones; y el último de 2018, con 228 observaciones. Todos ellos fueron el resultado de una encuesta realizada entre los dentistas de la clínica de Rui y cada conjunto de datos contiene 936 variables.

Dado que los conjuntos de datos tienen las mismas variables y no hay continuidad entre ellos, es decir, los pacientes de un año no se repiten en el mismo año o en los otros, o al menos no tenemos una manera de demostrarlo, hemos decidido considerar la totalidad de ellos como datos transversales y concatenarlos todos por fila. Esto resultó en un conjunto de datos con más de 2000 observaciones, más de 930 variables y sin distinción de tiempo, ya que no fue útil debido a la falta de continuidad y nuestros objetivos.

Volviendo a las variables, la gran mayoría son binarias, y como la cantidad de ellas era demasiado alta e identificamos una distinción temática entre algunas de ellas, hemos decidido dividirlas en 4 bloques teniendo en cuenta lo que queremos hacer con ellas y lo que necesitan nuestros objetivos. Estos bloques son las patologías sistémicas de los pacientes, sobre la intervención en sí misma, la medicación pre-intervención y la medicación post-intervención.

El primer bloque, las patologías sistémicas de los pacientes, contiene características suyas como el género, fechas de nacimiento, fechas de intervención, número de cigarrillos al día si son fumadores, colesterol, anemia, entre otros. El segundo bloque, la información sobre la intervención, contiene variables como la duración de esta, las herramientas usadas, el tipo de intervención, el material usado y cómo es, y mucho más. El tercer bloque y el último, que son medicación pre-intervención y post-intervención respectivamente, contienen las medicaciones que tomaban los pacientes por un lado antes de pasar por la intervención y, por otro lado, después de pasar por la intervención.

**::::::::::::::::::::::(Inglés)**

About the data used, to carry out our project, we were given some data from a dental clinic in Barcelona by the responsible of it, Rui Figueiredo. The data consists of three datasets: one from 2016; another one from 2017; and the last one from 2018. All of them were the result of a survey conducted among the dentists at Rui’s clinic, and each dataset contains 936 variables.

Since the datasets have the same variables and there is no continuity between them, we have decided to consider the entirety of it as cross-sectional data and to concatenate all of it by row. This resulted in a bigger dataset with over 2000 observations, over 930 variables and no distinction of time, as it was not useful.

Moreover, as the quantity of variables is too high, the vast majority are binary and we can identify a topic distinction between them, we have divided them into 4 blocks taking into account what we want to do with them and what our objectives need. These blocks are: the systemic pathologies of the patients, which contains characteristics such as gender, birthdate, number of cigarettes per day if they are smokers, if they drink alcohol or not, cholesterol levels, among others; the intervention details, which contains variables such as its duration, tools used and its characteristics, type of intervention, and much more; the medication pre-intervention, which is the medications that patients were taking before undergoing the intervention; and, finally, the medication post-intervention.

1. **Methodology (how):** (añadiría explicación sobre el tipo de metodología que hemos escogido, para qué ámbito está hecha en concreto, su objetivo y tal, he dejado un enlace abajo: 1 persona)

**::::::::::::::::::::::(Inglés)**

How did we organize our project? In order to do that we selected one of the most popular methodologies in the industry, which is CRISP-DM (it means Cross-Industry Standard Process for Data Mining), and we did some additions to it. With that methodology, the steps we followed and we are still following are the next ones:

1. Field of study:

* To read the guide about wisdom tooth removal that the dentist who sent us the data has provided us.
* Also to look up for information about the topic.

1. Data understanding:

* To make questions of interest and to determine some goals.

1. Data preparation:

* Data cleansing and division into 4 blocks.

1. Exploratory analysis:

* PCA: to look for determinant variables for perimplants disease.
* More options: exploration of variables like the intervention type to see the distribution.

1. Modeling:

* Model selection: random forest.

1. Evaluation:

* Evaluation methods: still pending.

1. Deployment:

* To write and correct the final report and presentation.

1. **Plan (when):** (enseñamos un diagrama de Gantt o algo similar con la planificación inventada o como sea de las tareas: 1 persona)
2. **Impact (whom):** (a quién le puede interesar o ayudar y lo qué podemos mejorar con nuestro estudio. Lo que dijo Rui, evitar operaciones innecesarias reduciendo el impacto económico y quizás ambiental cuando se hacen operaciones que terminan mal y se tienen que corregir: 1 persona)

Our project is conducted on common medical ground, being the three types of procedures, exposed earlier, a big percentage of interventions in the odontology field. Having a fruitful outcome, this research should prove interesting to many field specialists conducting these operations. Our aim with this project is to have a better understanding of the resources used in every one of these operations, meaning, time and staff. One of our major objectives is to improve the services offered by the professionals and the time used for every one of these procedures, diverting resources from those patients who require less clinical intervention, to the ones, due to a worse medical condition, who need it most.

We also aim at improving the understanding of the different types of patients, and the complications that might emerge because of their health conditions. Finding correlations between common health issues, medication used before, during and after the interventions with favorable or unfavorable outcomes of the procedure might enlighten the reasons for possible postoperative complications. To sum up, recognizing individual characteristics of the patients will offer an alternative and more personalized perspective on medical interventions.

1. **Referencias bibliográficas:**

<https://www.ibm.com/docs/en/spss-modeler/saas?topic=dm-crisp-help-overview>