

# Análise e Previsão de Alterações no Funcionamento da Rede: Relatórios de Situação e Planos da Semana de PFACD

André Plancha  
105289, CDC2  
Andre\_Plancha@iscte-iul.pt

João Francisco Botas  
104782, CDC1  
Joao\_Botas@iscte-iul.pt

Marco Esperança  
110451, CDC1  
marco\_esperanca@iscte-iul.pt

Allan Kardec Rodrigues  
103380, CDC1  
aksrs@iscte-iul.pt

Diogo Alexandre Alonso de Freitas  
104841, CDC1  
daafs@iscte-iul.pt

## Índice

II. Plano semanal: 22 fevereiro (Semana 2 - 3) .....	2
Informação da Reunião .....	2
Feedback geral e recomendações .....	2
Informações e Conclusões .....	2
Plano semanal .....	2
Organizar e sincronizar áreas de trabalho .....	2
Pedir dados adicionais .....	2
Analisar a tabela de indicadores de continuidade de serviço .....	3
Começar e adiantar o 1º <i>sprint</i> .....	3
I. Relatório de Situação: 20 fevereiro (Semana 2) .....	3
Planeamento das entregas .....	3
Problemas que bloquearam o trabalho durante a semana .....	3
Riscos que podem prever para o desenvolvimento do projeto e planos de contingência .....	4
Progressos feitos .....	4
Descrições dos dados e análise da entidade .....	4
Análise dos dados e Visualizações .....	4
Próxima semana .....	5

## II. Plano semanal: 22 fevereiro (Semana 2 - 3)

### Informação da Reunião

A reunião foi feita ontem (21/2), e todos os membros do grupo estiveram presentes. Os porta-vozes de Diogo Alexandre Alonso de Freitas e Marco Esperança, e próxima semana será João Francisco Botas o porta-voz, desta vez eleito pelo professor, posteriormente à reunião.

### Feedback geral e recomendações

- O relatório semanal entregue foi demasiado denso, e recomendou fazer por tópicos;
- Temos de ser mais específicos com as datas de entrega e conteúdos dos *sprints*;
- Embora a falta de dados possa ser problemática, não impossibilita o tema escolhido, sendo que podemos fazer uma análise geográfica e setorial;
- Incluir o tema mais claro nos entregáveis.

### Informações e Conclusões

- A reunião com a entidade será daqui a 2 semanas (semana 4);
- A entidade não irá disponibilizar dados pré 2020 até abril, e recomenda que usemos os dados de indicadores gerais por concelho pois incluem dados desde 2014;
- O professor não obteve resposta do IPMA.
- O professor pode disponibilizar dados sobre incêndios urbanos, providenciados pela ANAPEC.
- O tema escolhido será sobre as alterações assinaláveis no funcionamento da rede durante os períodos de vigência dos avisos de perturbações atmosféricas (Projeto 3), o qual nós decidimos como objetivo principal a análise e previsão no funcionamento da rede (potencialmente de forma contínua e de forma difusa).
- Até abril, devemos já ter um primeiro modelo e avaliação supervisionada, o qual estará presente no 3º *sprint*;
- Depois de abril, será feita um segundo ciclo do CRISP-DM, de forma a rever as conclusões e suposições nas várias fases iniciais do processo, e encontrar outras formas de limpeza e preparação dos dados.

### Plano semanal

Esclarecendo as dúvidas que tínhamos sobre o tema, esta semana iremos começar em concreto a análise à volta dele. Especificamente, iremos:

- Organizar e sincronizar áreas de trabalho
- Pedir os dados adicionais recomendados e necessários
- Analisar a tabela de indicadores de continuidade de serviço
- Começar e adiantar o 1º *sprint*

### Organizar e sincronizar áreas de trabalho

Neste momento, estamos a usar várias ferramentas para organizar a informação colecionada e sincronizar trabalho. O nosso plano é completar a sincronização das áreas de trabalho para poder trabalhar em síncrono. Mais especificamente, vamos usar:

- O *git* e *github* para organizar documentos encontrados úteis (como o manual de ligações da e-redes), e partilhar e sincronizar *scripts* e *notebooks*;
- *Typst Web Application*<sup>1</sup> + *PowerPoint*, para sincronizar a escrita dos relatórios e dos *sprints*;
- *pixi*<sup>2</sup>, para poder sincronizar bibliotecas de python e outras ferramentas, usando *Conda*, de forma a ter um ambiente reproduzível;

entre outras que possivelmente surgem-se úteis durante a semana

### Pedir dados adicionais

Como recomendado, iremos pedir vários dados adicionais para a elaboração do projeto. Em princípio, iremos pedir:

---

<sup>1</sup><https://typst.app/>

<sup>2</sup><https://pixi.sh/latest/>

- Dados meteorológicos (temperatura, precipitação, humidade, intensidade e direção do vento, pressão atmosférica, trovoadas, descargas elétricas), geofísicos (sismos), e potencialmente climáticos (ondas de calor, monitorização da seca) ao IPMA;
- Dados de incêndios urbanos ao professor (provenientes da ANEPC, disponibilizados pelo professor)

Adicionalmente, iremos pesquisar sobre dados demográficos de Portugal que nos possam ser úteis. Estes provavelmente estarão disponíveis na plataforma PORDATA.

## Analisar a tabela de indicadores de continuidade de serviço

TODO

### Começar e adiantar o 1º *sprint*

O primeiro *sprint* será principalmente sobre BU. Planeamos que o *sprint* terá uma descrição da entidade, um entendimento detalhado do problema que vamos analisar, uma revisão da literatura relevante, e uma descrição dos dados que nos são disponíveis que nos possam ser úteis, ambos da entidade e de terceiros<sup>3</sup>. Planeamos que, para terça-feira (27/2) teremos já um primeiro rascunho (pelo menos da organização do documento), de forma a o professor poder dar *feedback*.

## I. Relatório de Situação: 20 fevereiro (Semana 2)

Esta semana passada, devido à natureza do projeto, começámos a fazer a fase inicial do projeto. Isto significa que fizemos uma investigação superficial à entidade dos dados, uma descrição dos dados que nos são disponíveis no momento e das variáveis de tais, e começamos a fazer algumas visualizações iniciais. Adicionalmente, requisitamos os dados adicionais ao professor. Estes progressos foram de grande parte feitos de forma independente, devido à natureza do realizado, às limitações descritas a seguir, e ao Carnaval.

Devido aos dados limitados que nós temos acessos, e devido à falta de comunicação que tivemos até ao momento, as análises feitas foram substancialmente limitadas ao conhecimento e informação disponível pelos membros do grupo. Esperamos que a reunião presencial esclareça várias dúvidas que o grupo tem, diretamente ou indiretamente.

### Planeamento das entregas

Antes da pausa letiva da Páscoa, planeamos já ter concluído, ou quase concluído, até à fase de Feature Engineering; ou seja, planeamos ter finalizadas as fases de entendimento do negócio, de reconhecimento dos dados e de transformação e extração de variáveis. Como estimativa planeamos entregar:

- O primeiro *sprint* na semana 3 (26/2 - 3/3, próxima semana),
- O segundo *sprint* na semana 6 (18/3 - 24/3),
- O terceiro *sprint* na semana 9 (8/4 - 14/4).

Estas estimativas são completamente dependentes da disponibilização de dados adicionais pelo distribuidor, e não são definitivas. Como solicitado, cada *sprint* será alinhada com as primeiras fases do CRISP-DM. Para esse fim, os *sprints* serão sobre Business Understanding (BU), Data Understanding (DU), e Data Preparation (DP). Esperamos que mais promenores sobre o que cada *sprint* precisará fique mais claro na reunião de quarta e ao longo do tempo.

### Problemas que bloquearam o trabalho durante a semana

Apesar de existirem vários dados presentes na fonte disponibilizada pela E-REDES<sup>4</sup>, estes são limitados devido às suas datas, que a maioria é de 2021 para a frente, tornando-se difícil fazer comparações com a época do covid-19 e da crise económica de 2008 (entre outras).

Na mesma fonte, existem também tabelas com variáveis “codificadas”, tornando-se difícil a compreensão de muitos dados apresentados. Isto nota-se especialmente nas tabelas de indicadores de continuidade de serviço<sup>5</sup>

Também se torna difícil relacionar os dados que temos acesso atualmente com dados meteorológicos e afim, sendo que não encontramos nenhuns dados públicos históricos, inclusive no website do IPMA. Os dados que encontramos

<sup>3</sup>Dependente da disponibilização dos dados até a data planeada de entrega do *sprint* pelos terceiros

<sup>4</sup><https://e-redes.opendatasoft.com/explore/>

<sup>5</sup>Exemplo: <https://e-redes.opendatasoft.com/explore/dataset/12-continuidade-de-servico-indicadores-gerais-de-continuidade-de-servico/table/>

úteis sobre o tópico são sobre séries longas<sup>6</sup>, e a maioria são de 2020 para traz, por distrito (não estando todos) e a intervalos mensais, apenas tendo acesso a conjunto de indicadores sobre a temperatura, precipitação e pressão (excepto Lisboa). A utilidade destes dados é volátil e debatível.

## Riscos que podem prever para o desenvolvimento do projeto e planos de contingência

Neste momento, devido à incerteza do que o objetivo principal do projeto é, potenciais riscos e os seus planos de contingência tornam-se difícil de calcular. Devido a tal, ainda gostaríamos de ter mais informação adicional e trocar impressões com o professor antes de responder a esta questão. O nosso principal risco no momento é na disponibilização dos dados adicionais. Isto deve-se ao facto da nossa ideia inicial ser estudar as crises pandémica e económica mas não termos dados de antes de 2020 em bases de dados que achámos relevante estudar. Se a disponibilização não acontecer num futuro próximo, será provavelmente necessário uma mudança radical do plano atual.

## Progressos feitos

### Descrições dos dados e análise da entidade

Com isto que dissemos das informações e variáveis existentes na base de dados, podemos referir que nesta semana procurámos fazer uma pequena descrição de cada uma das 21 bases de dados e ver possíveis lacunas que estas continham, verificando limites de datas nos dados (concluindo que não são de todo estandardizados entre as 21), e verificar a importância que poderão ter para o projeto. Esta descrição está disponível no primeiro rascunho do primeiro *sprint*.

As descrições não estão completas, sendo que neste momento ainda esperamos pelos dados adicionais requisitados.

Na execução disto, foram também descobertas diversas informações sobre a E-REDES, o que nos levou a encontrar diversos documentos e informações relacionadas com o projeto. Estes documentos e informações foram partilhados entre o grupo, e no futuro é planeado sincronizar e disponibilizar estas. Diversas informações mais gerais foram também anotadas e estão disponíveis no mesmo rascunho.

### Análise dos dados e Visualizações

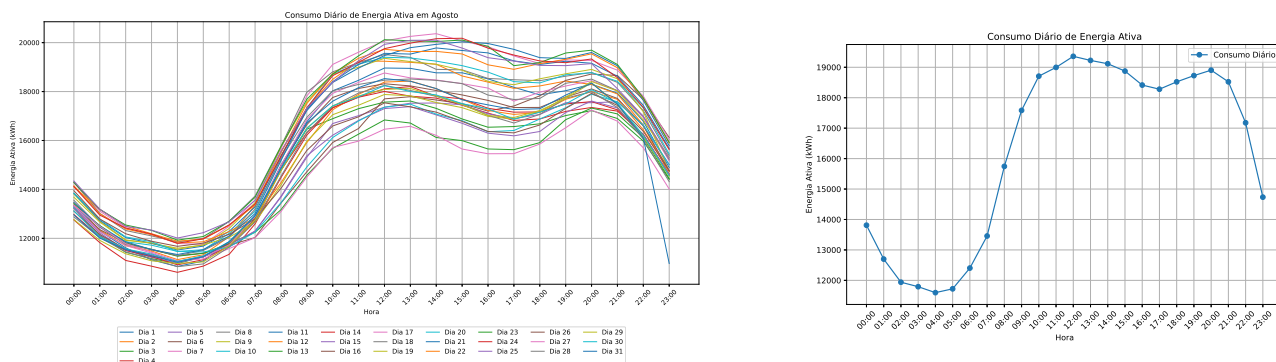


Figura 1: Consumo energético diário no mês de agosto em Cascais, por dia

Na Figura 1 podemos visualizar o consumo, em 2023, em Cascais. À esquerda, podemos realizar uma breve comparação de todos os dias, enquanto o gráfico da direita demonstra como foi o consumo só do dia 9 no mesmo local. Para, uma evolução ao longo do mês, abaixo será possível visualizar o respetivo consumo de cascais ao longo de todo o mês de agosto (é importante referir que a *label x* encontra-se codificada em horas, desde o dia 1/00:00h até ao dia 31/23:00h).

<sup>6</sup><https://www.ipma.pt/pt/oclima/series.longas/list.jsp>

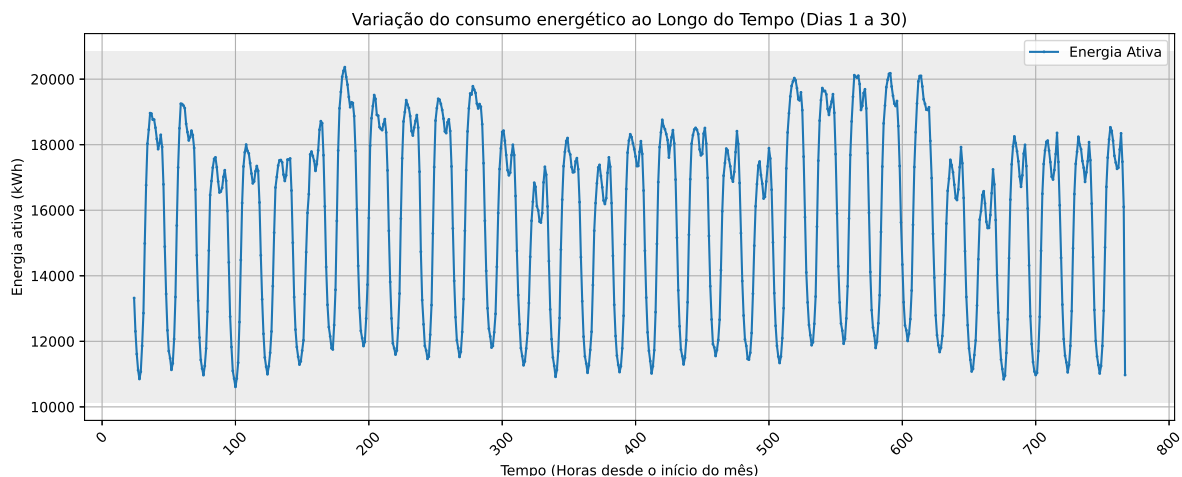


Figura 2: Consumo energético no mês de agosto em Cascais, contínuo

A Figura 2 demonstra que o consumo energético é, de certa forma, cíclico, isto é, usando os gráficos anteriores como base, em praticamente todos os dias, existe um maior consumo às 14h e um menor consumo às 23h.

**NOTA:** Apenas possuímos estes dados na época de Agosto de 2023 e, apesar desta limitação, gostaríamos de realizar os mesmo gráficos para uma zona turística/férias a fim de comparar o consumo.

### Próxima semana

Para avançar neste projeto é importantíssimo ter a reunião de amanhã, pois o grupo acredita que poderá esclarecer as várias questões que possui, desde a obtenção total dos dados, como a escolha do tema em si. Esperamos ter o primeiro *sprint* acabado na próxima semana, esse em princípio será o foco da próxima semana. Isto será dependente do resultado da reunião de amanhã, da data da reunião com a entidade, e da disponibilização dos dados adicionais.

Esperamos, depois da reunião, conseguir arranjar um objetivo claro e fixo, a fim de, finalmente, criar o foco essencial para prosseguir no trabalho de forma eficaz e produtiva.