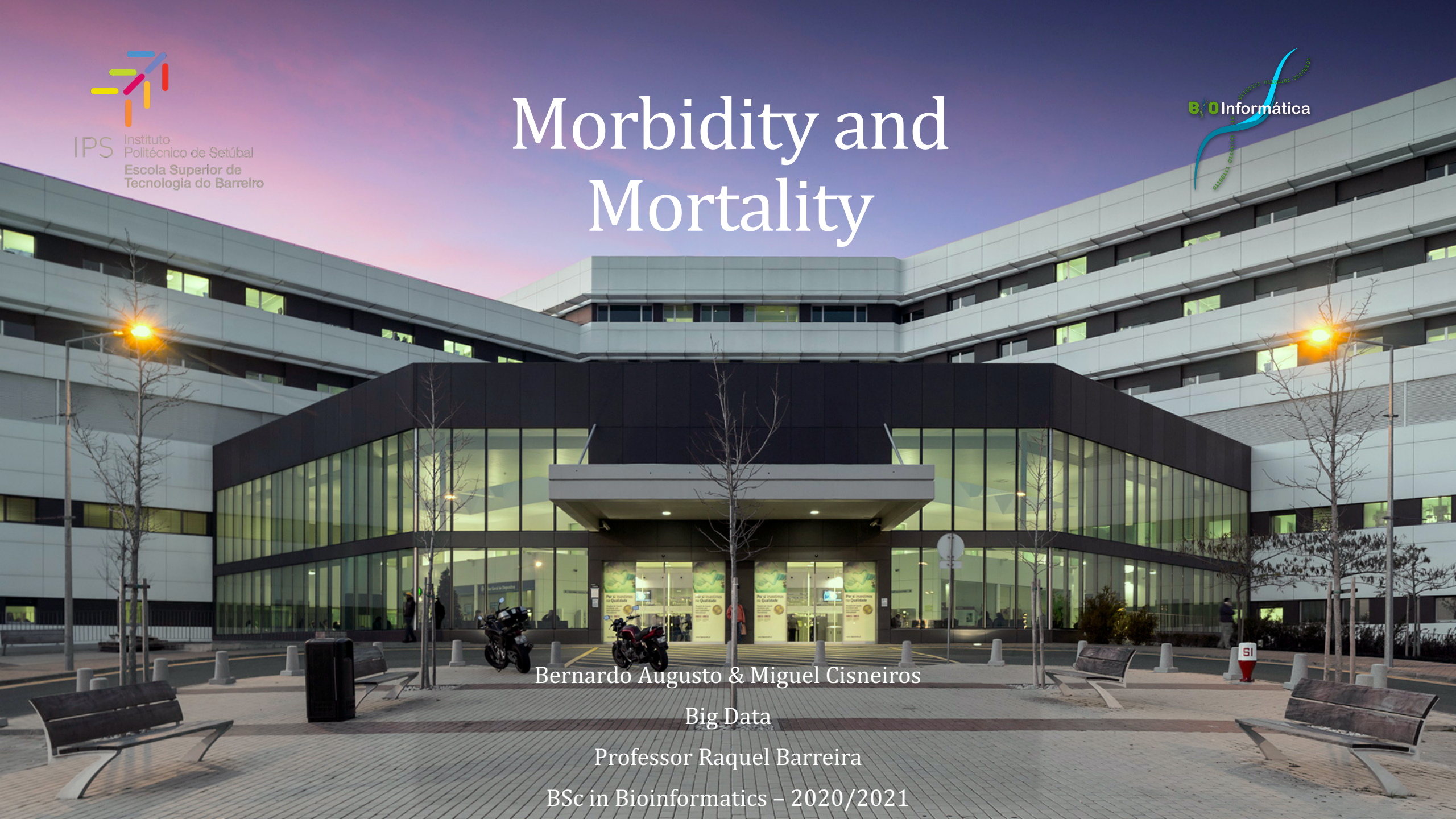


Morbidity and Mortality

Bernardo Augusto & Miguel Cisneiros

Big Data

Professor Raquel Barreira

BSc in Bioinformatics – 2020/2021

Overview of the task

The goal: To explore a document-oriented database in MongoDB

Methodology

1. Check open-access database sources
2. Select a national database
3. Upload and work it on MongoDB (with queries and aggregations)

CASCAIS DATA › Hospital Cascais Dr. José de Almeida › mongo DB

The document dataset

- This dataset is related to the monthly evolution of hospitalization, outpatient and death episodes by main diagnosis category, between January 2016 to September 2018 at the Cascais Hospital - Dr. José de Almeida.
- It is provided in a JSON file (JavaScript Object Notation) and contains the following attributes:

- datasetid – identification of the dataset
- recordid – identification of the row
- fields/obitos – number of deaths ocurred
- fields/dias_internamento – number of days of hospitalization
- fields/faixa_etaria – age range of the patient
- fields/periodo – month and year of the episode
- fields/desc_capitulo – main diagnosis category

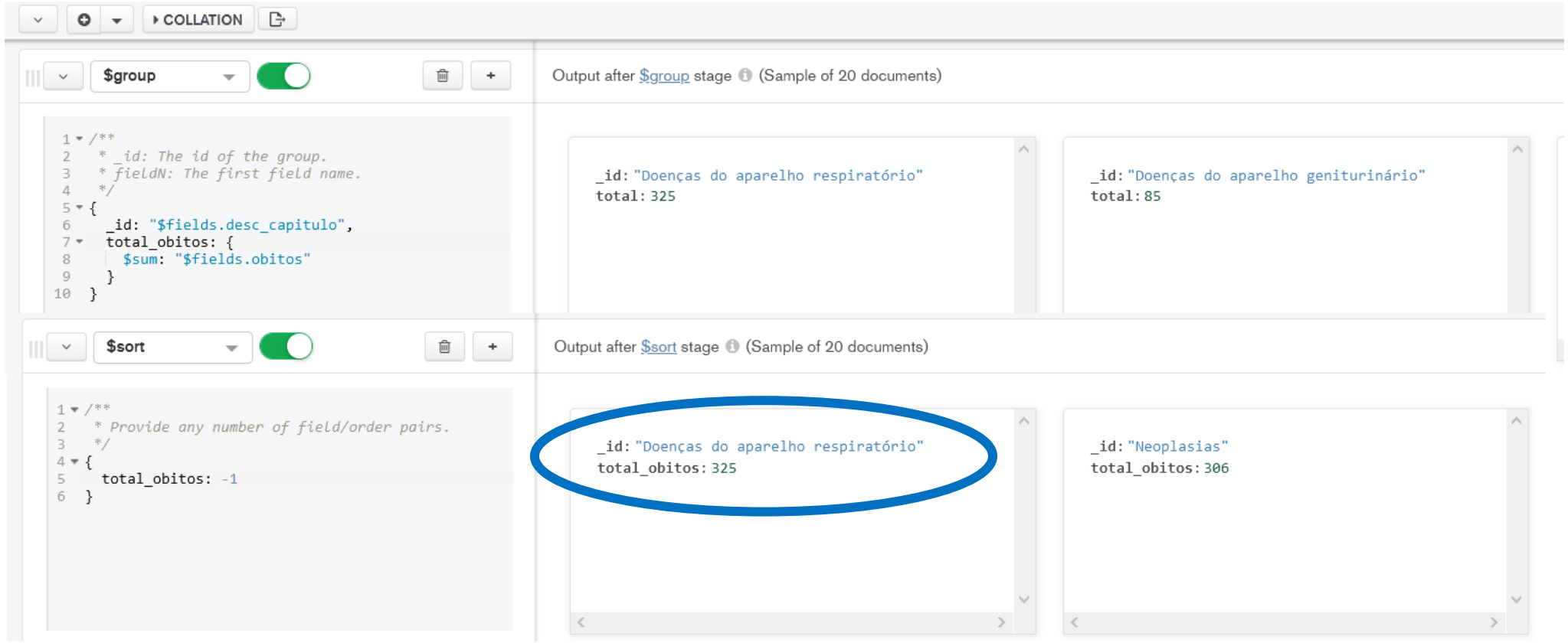
- fields/sexo – gender of the patient
- fields/regi_es - health region of the hospital
- fields/internamentos – number of hospitalized patients
- fields/instituicao – name of the hospital
- fields/cod_capitulo – code of the main diagnosis category
- fields/ambulatorio – number of outpatients
- record_timestamp – record date and time

cascais.cascais

COLLECTION SIZE: 2.39MB TOTAL DOCUMENTS: 5004 INDEXES TOTAL SIZE: 72KB

Query 1

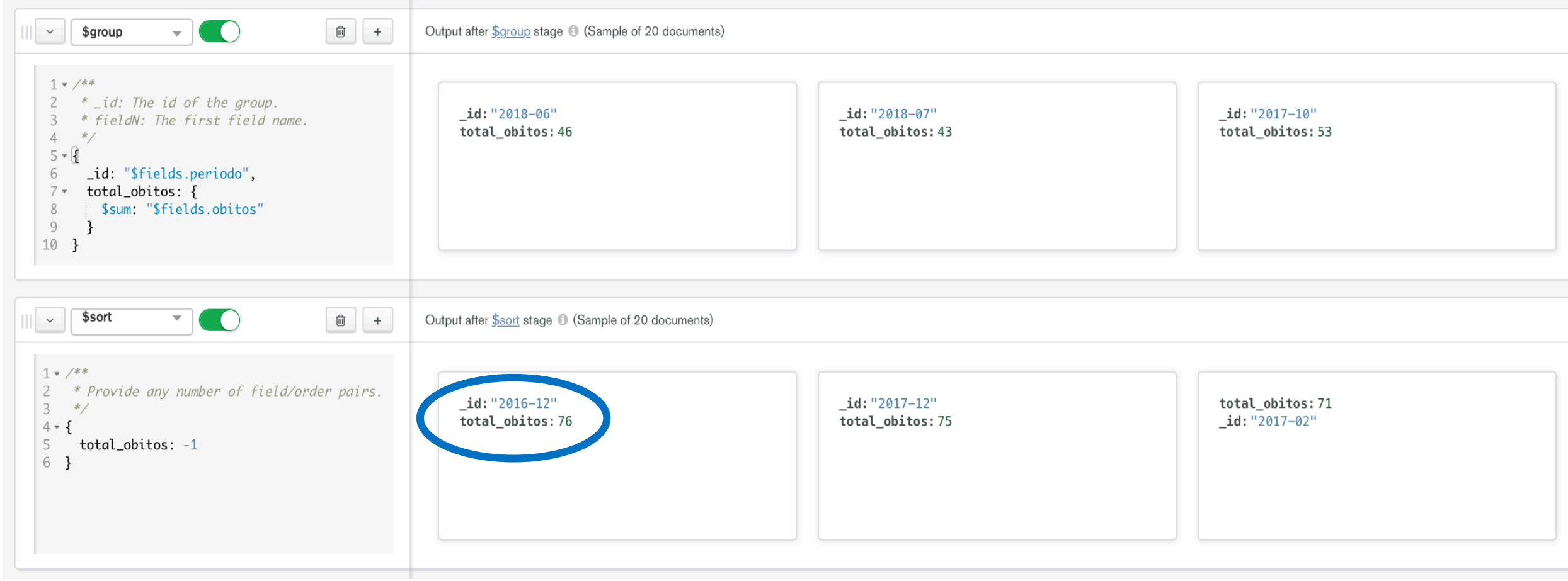
Q: What was the main cause of death in the timeframe of the dataset?



A: Diseases Of The Respiratory System

Query 2

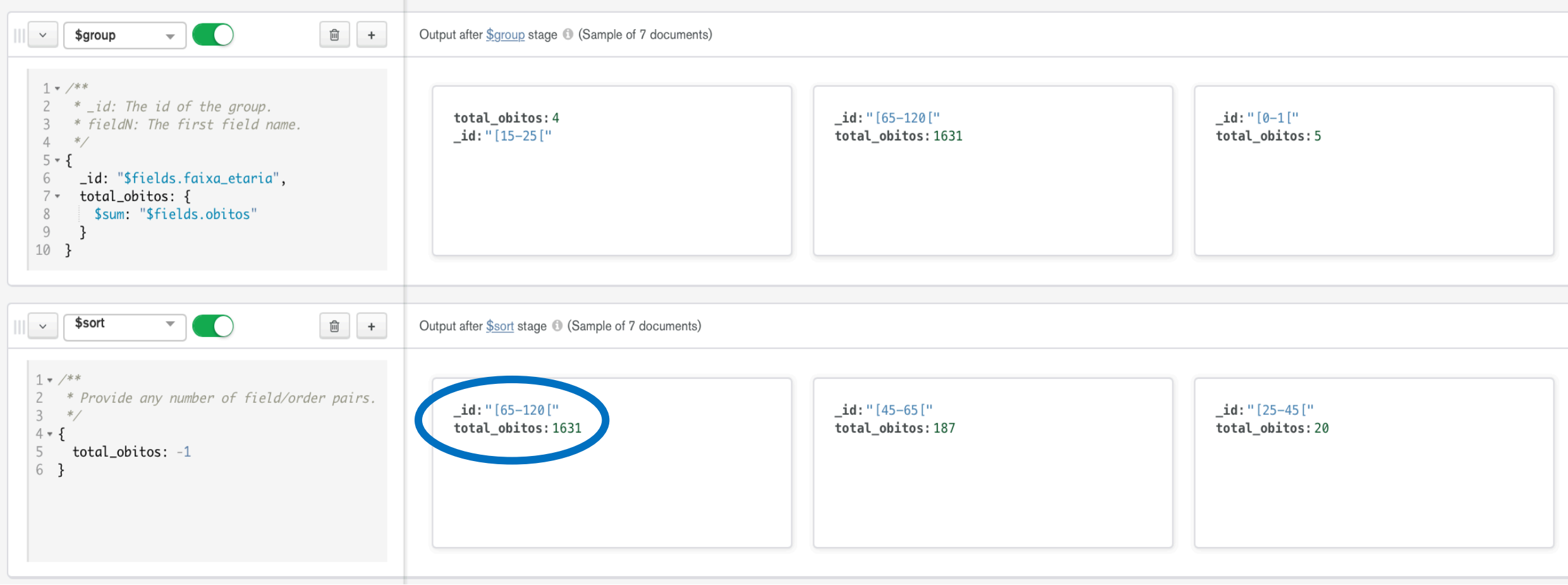
Q: What was the month with more deaths?



A: December 2016 with a total of 76 deaths

Query 3

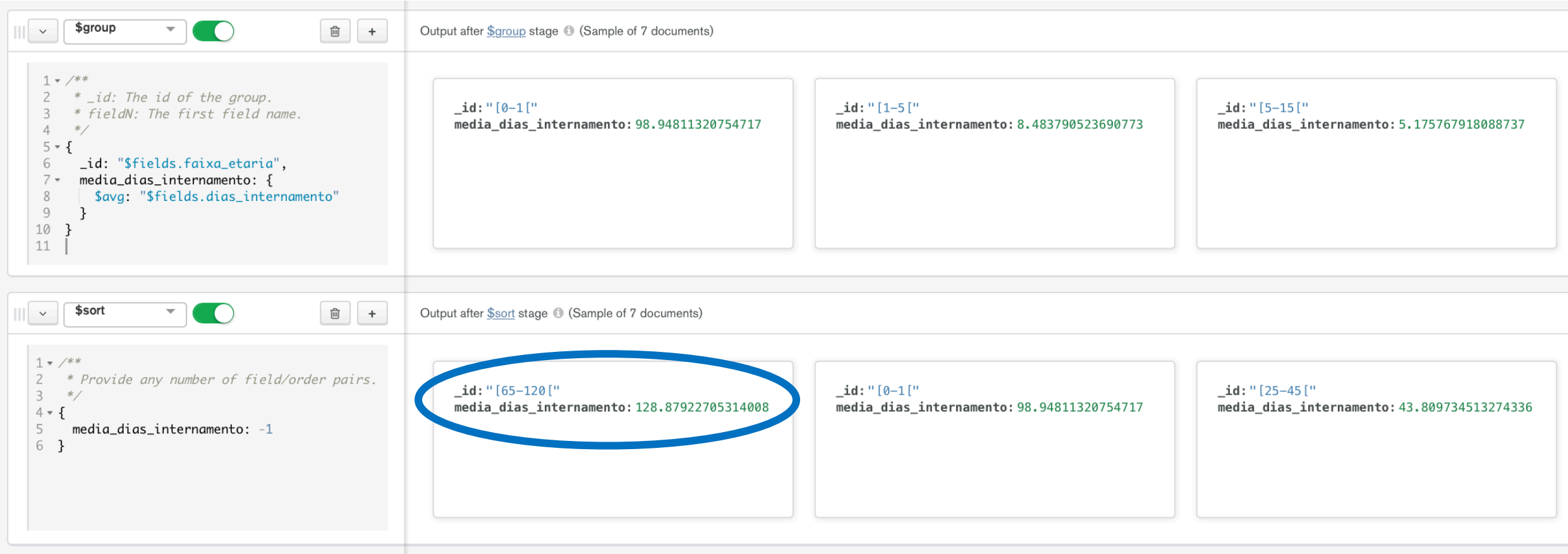
Q: What was the age group with more deaths in the timeframe of the dataset?



A: People from 65 to 120 years old were the ones with more deaths

Query 4

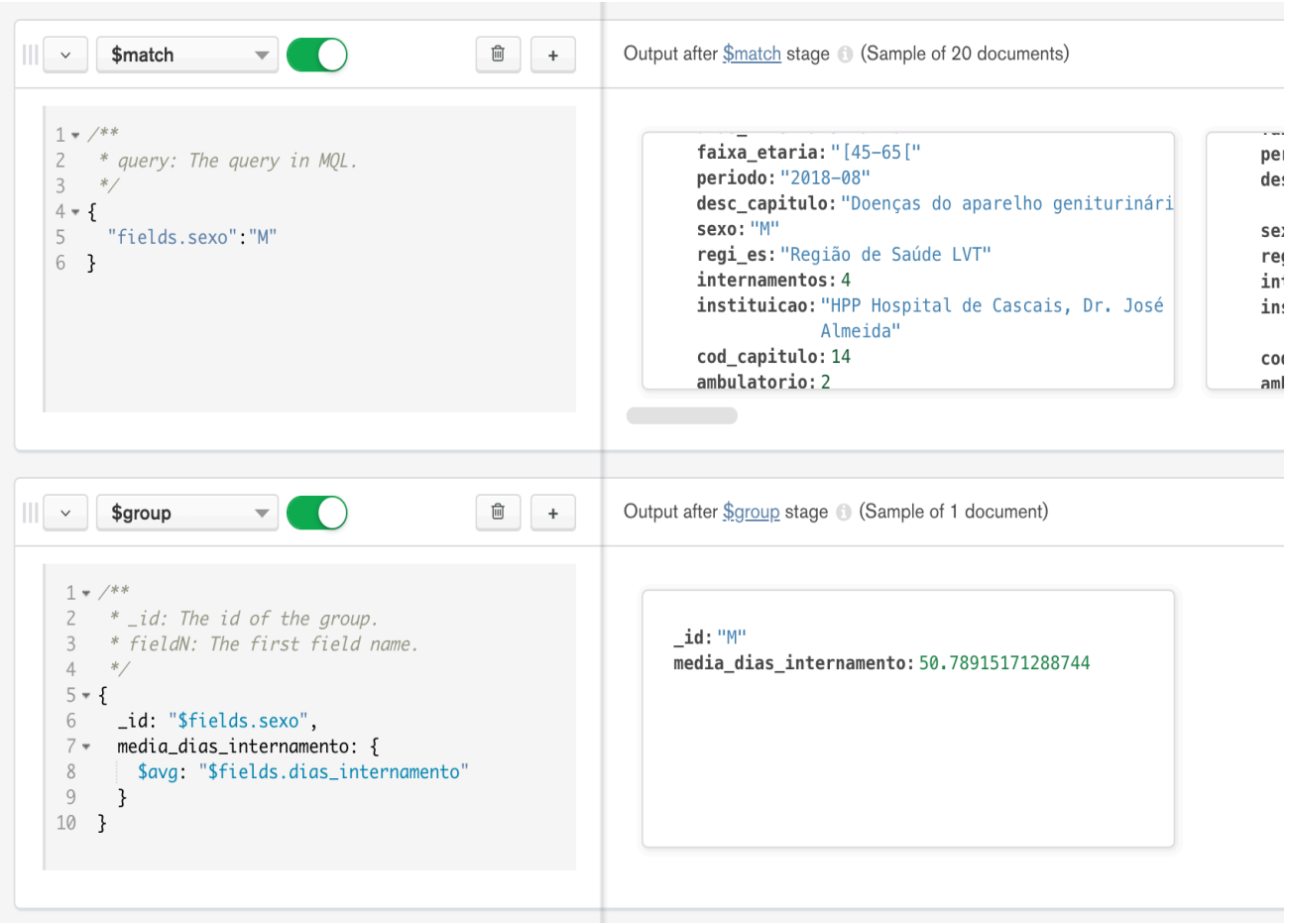
Q: In average, what was the age group with more days of hospitalization?

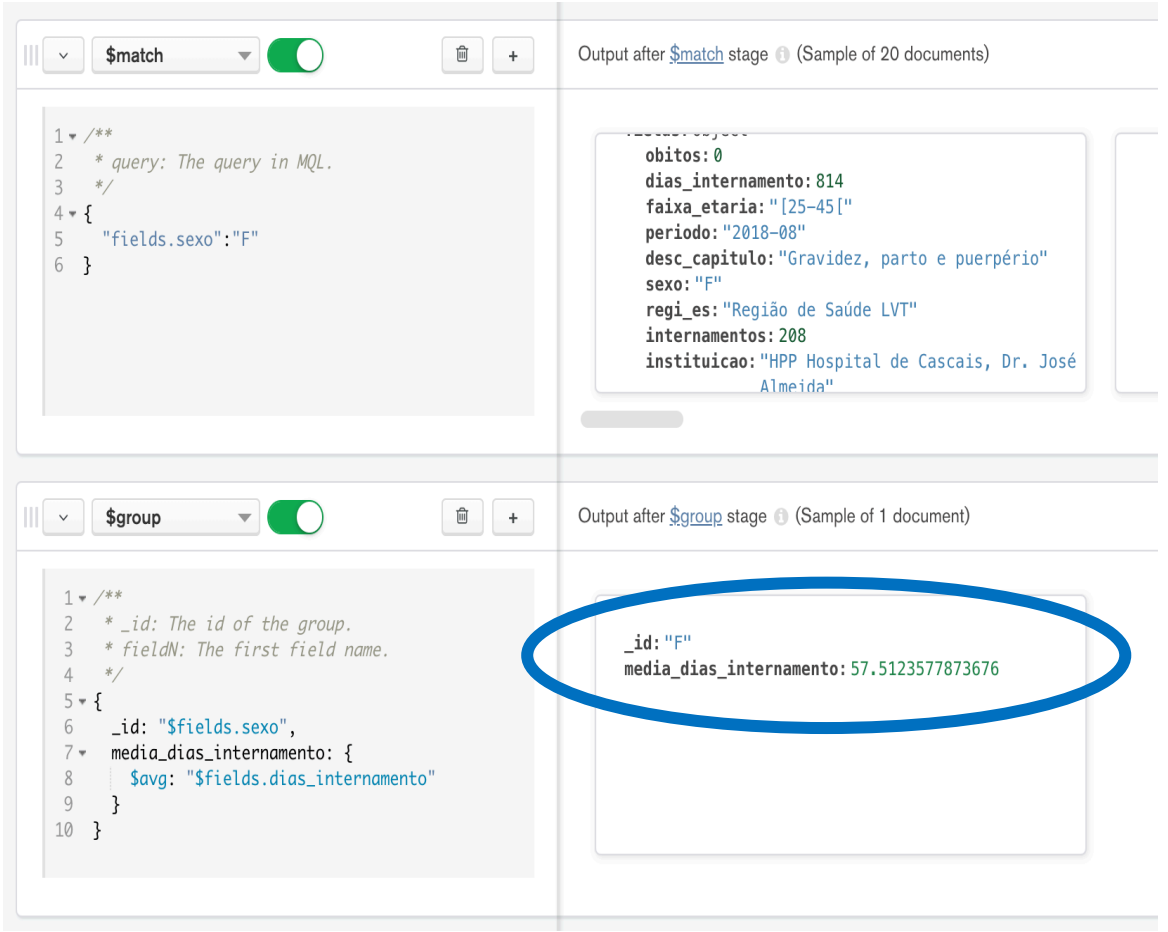


A: People from 65 to 120 years old were the ones with more days of hospitalization

Query 5

Q: In average, what was the gender with more days of hospitalization?





A: Women had, in average, more days of hospitalization than men

References

- MongoDB. (n.d.). *Get started with MongoDB — MongoDB Documentation*. MongoDB | Documentation. https://docs.mongodb.com
- MongoDB. (2019, June 13). *MongoDB Atlas - Aggregation Pipeline Builder Tutorial* [Video]. YouTube. https://www.youtube.com/watch?v=0MZFTiKIPnU
- Barreira, R. (2020). *06 - MongoDB* [Slides]. Moodle | IPS. https://moodle.ips.pt/2021/pluginfile.php/197606/mod_resource/content/1/BD2021_06.pdf
- Database https://data.cascais.pt/pt-pt/node/254

Thank you


BIO Informática
IPS Instituto Politécnico de Setúbal
Escola Superior de Tecnologia do Barreiro