Portfolio 1

Open Data Science

# Titanic datasæt

*Opg. 1* Åben filen i en tekst-editor og se på indholdet.

*Datatyper:*

**Integers** for rejseklasse, antal af søskende/ægtefælle, antal af forældre/børn ombord. Integers bruges også til alder, men der er enkelte mennesker med sin alder i **float**.

Derudover anvendes der 1 eller 0 som betegner om personen overlevede eller ikke.

Navne angives i **strings** ligeledes køn.

Kolonnen der beskriver *fare* betegnes ved floats, og nogle steder integers.

Nogle fares er betegnet som 0 som kunne indikere manglende data.

*Opg. 3*

Der var i alt 342 overlevende fordelt på alle klasser.

Gennemsnits alderen er 29.5 og median ligger på 28 år gammel.

Ældste passager var 80 år gammel, mens den yngste var knap fire måneder

*Opg. 5*

Pivot-table øvelse viser at rejseklasse 2 havde flest omkomne, efterfulgt af rejseklasse 3, med rejseklasse 1 med flest overlevende passagere.