Ikea:

1. Clean the dataset/Check data types
2. Perform PCA for dimensionality reduction (explain 90-95% of total variance) - exclude categorical data / textual + normalize/scale data
3. Apply Elbow method to select optimum number of clusters - would turn out to be 3
4. Apply K-Means (feed the results from PCA to model)
5. Append column clusters to original df
6. Append a column has_ikea to original df
7. Explore clusters - you will probably find them separated into small medium and large municipalities
8. Analyse based on for example: population, sales index, revenue.. exclude where ikea already exists, close to existing ikea
9. Suggest top 3 candidates to open new ikea store.

Titanic: (hint: the biggest task here is preparing the data)

1. Explore the dataset (understand the features and make some basic assumptions)
2. Think what features would have an effect on survival rate (you would probably want to drop the ones that don't). This may be risky though, so you could opt for performing a Correlation analysis to see which features are most correlated to survival. Then you are not relying so much on guessing.
3. pre-processing: Missing data, converting features (to numeric), creating categories, creating new features
4. Fit classifier / evaluate performance using k-fold cross validation
5. Analyse