**INTD 8065 Data Analysis for Cancer Research**

**Homework #2**

Use dataset VA in library MASS. This dataset contains the following variables:

stime: survival or follow-up time in days.

status: dead (1) or censored (0).

treat: treatment: standard (1) or test (2).

age: patient's age in years.

Karn: Karnofsky score of patient's performance on a scale of 0 to 100. 10-30 completely

hospitalized, 40-60 partial confinement, 70-90 able to care for self.

diag.time: times since diagnosis in months at entry to trial.

cell: one of four cell types; squamous (1), small cell (2), adeno (3) or large cell (4).

prior: prior therapy; No (0), yes (10).

1. Only nine of the 137 patients survived at the end of the study. Using dplyr, filter the dataset so it only contains patients who died.
2. Produce a barplot for the cell types of tumors.
3. Are the different cell types equally represented in treatment groups?