R Programming Formative Assessment Instructions

**Background**

Heart and circulatory diseases cause around a quarter of all deaths in the UK, according to the British Heart Foundation. In fact, it is estimated that there is nearly 1 heart attack hospital admission every 5 minutes in the UK.

In this analysis, you will be exploring the **Heart Disease Statistics by Health Board** data sets from Public Health Scotland. This is the data that we worked with in the Week 2 tutorial.

**Data**

Three data files are needed for your analysis:

1. Heart Disease Activity By Health Board:
   1. <https://www.opendata.nhs.scot/dataset/scottish-heart-disease-statistics/resource/748e2065-b447-4b75-99bd-f17f26f3eaef>
2. Heart Disease Mortality By Health Board:
   1. <https://www.opendata.nhs.scot/dataset/scottish-heart-disease-statistics/resource/dc0512a8-eb49-43b9-84f1-17ef95365d57>
3. Health Board look up:
   1. <https://www.opendata.nhs.scot/dataset/9f942fdb-e59e-44f5-b534-d6e17229cc7b/resource/652ff726-e676-4a20-abda-435b98dd7bdc>

You can either load the data from the website using the URL or download the data onto your device and load it into R. Be sure to include clear documentation in your code of which approach you take.

**Assignment**

The purpose of this assignment is to provide an optional formative assessment. You will not receive a mark for this assignment, nor are you required to complete it. It is intended to provide an additional learning opportunity for you to get feedback from the teaching team. The assignment is quite similar, though shorter, than the marked R Programming assignment at the end of the course.

You can use any R packages of your choosing to create this report; you are not restricted to those discussed in the course thus far or those included in the helper code.

To help you get started, we have included an example of a basic script which shows some code for reading in the data plus some basic data wrangling (formative\_helper.Rmd). You will notice, however, that there are not any comments in this code.

**Question**

In your analysis, explore **one** the following questions:

1. What type of heart disease is the most common according to age?
2. Which diagnosis leads to the highest rate of death and for which demographic (the data include age and sex)?

We are not looking for one specific analysis. As you work with the data, feel free to make decisions on whether you want to focus on a specific year or location(s) if that makes your message clearer and analysis more sound. You can interpret the question however you like (e.g., looking at a specific age range, all age ranges, some compared to others, etc.).

**Post your code, with comments, and figure to the Formative Assessment Discussion Board to receive feedback from the teaching team.** There is no deadline for this formative assessment.

**Tips**

* Be clear in how you are specifically operationalizing the question.
* Watch out for the aggregate level variables in the data set
* Look carefully at the helper code and make sure you are filtering the data as you are intending to