HW\_2

library(tidyverse)

library(stringr)  
got\_chars\_full <- read\_csv("character-predictions\_pose.csv")

**Отбор релевантных переменных в таблицу got\_chars**

got\_chars <- got\_chars\_full %>%   
 select("name", "title", "male", "culture", "dateOfBirth", "DateoFdeath", "house", "spouse", "age")  
rm(got\_chars\_full)

**Переименование значений переменной.** Насколько я понял, мы создаем список для индексирования. Не стал этого делать, поскольку сам этим приемом обычно не пользуюсь.

got\_houses <- got\_chars %>%   
 filter(house %in% c("Night's Watch", "House Frey", "House Stark", "House Targaryen", "House Lannister", "House Greyjoy")) %>%   
 select(name, house) %>%   
 mutate(house =   
 str\_replace(house, ".\* ", paste0(str\_extract(house, "^."), "\_")) %>%  
 tolower())

*Типичный пример, когда после того, как заморочился с регулярными выражениями, понимаешь, что проще было напечатать все, что нужно)))*

**Создание таблицы с информацией по супругам**

# Создаю вспомогательную таблицу. На всякий случай, предварительно фильтрую пропущенные наблюдения. Наверное, можно было и с ними, но не хотелось проверять.  
assist <- got\_chars %>%   
 filter(!is.na(spouse)) %>%   
 mutate(spouse\_name = name,  
 spouse\_title = title,  
 spouse\_age = age) %>%   
 select("spouse\_name", "spouse\_title", "spouse\_age")  
  
# Объединяю основную и вспомогательную таблицы, отбираю нужные переменные. Переименовываю переменную spouse, поскольку у ключа сохраняется то имя, которое было у нее в основном датасете.  
got\_spouses <- got\_chars %>%   
 filter(!is.na(spouse)) %>%   
 left\_join(assist, by = c("spouse" = "spouse\_name")) %>%   
 select("name", "title", "male", "age", "spouse", "spouse\_title", "spouse\_age") %>%   
 rename(spouse\_name = spouse)

**Удаление переменной**

got\_spouses <- got\_spouses %>%   
 select(-male)

**Выбор нужных переменных и изменение их расположения.**

got\_spouses <-got\_spouses %>%   
 select( "name", "spouse\_name", "title",  
 "spouse\_title", "age", "spouse\_age") %>%   
 mutate(age\_diff = abs(age-spouse\_age))