##title: “Jeffrey_exercise4_hw1” ##author: “Jack Jeffrey” ##date: “2024-09-09” ##output: html_document

Exercise 4

rm(list=ls(all=TRUE)) #cleaned up workspace

cat(“\014”) #cleaned up console

1.

install.packages(“tidyverse”) library(tidyverse) # instealled tidyverse and dplyr

library(nycflights13) head(flights) # Check the dataset

1.1 find all March flights

str(flights) march_flights <- flights %>% filter(month == 3) head(march_flights) # successfully located March flights

1.2 new variable using mutate

flights_with_dates <- flights %>% mutate(date = make_date(year, month, day)) head(flights_with_dates) # new date variable created using mutate

1.3 change column name

flights_rename <- flights %>% rename(tail_number = tailnum) view(flights_rename) # created new dataset with corrected column name

1.4 group flights by their origin

flights_origin <- flights %>% group_by(origin) str(flights_origin) group_vars(flights_origin) # flights grouped by origin

1.5 counting JFK departures 2013

jfk_departures <- flights %>% filter(origin == “JFK”, year == 2013, month == 12, day ==31) %>% summarize(count = n()) jfk_departures # 283 departed flights from JFK on 12/31 2013

1.6 calculating average delays at JFK

jfk_delay_average <- flights %>% # created the object for JFK flight delays filter(origin == “JFK”) %>% # selected flighst from JFK summarize(avg_delay_hours = mean(dep_delay, na.rm = TRUE) /60) # calculated the average delay in hours and removed na jfk_delay_average # average delay = 0.202 hours or about 20 minutes