

Domáca úloha 1 - Moderná aplikovaná regresia

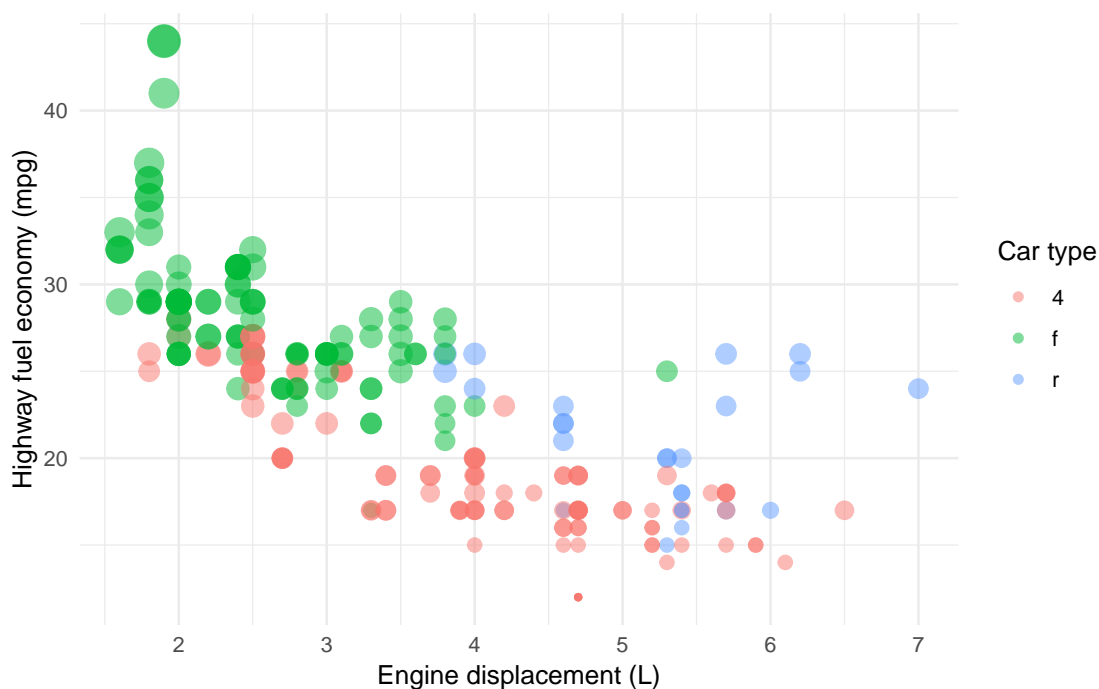
Prosím odovzdať do najbližších cvičení.

1 (70b)

Majme dataset `divusa` z knižnice `faraway` spravte numerické a grafické zhrnutie dát (tak ako na cvičeniach), poriadne okomentujte vaše zistenia a vlastnosti, ktoré sa vám budú zdať zaujímavé pre budúcu analýzu. Výsledky na konci zhrňte (sformulujte) do jedného záveru aby ste sa k tomu mohli kedykoľvek vrátiť. Riešenie by malo obsahovať približne tieto body:

- Načítanie a oboznámenie sa s problematikou (aké máme regresory, čo z toho je odozva)
- Bližší náhľad na dáta a popisná štatistika
- Oprava a Transformácia (ak je nutná)
- Prvotná vizualizácia (klasike a krížové zobrazenia odozvy vs. prediktor(i))
- Hlbšia vizualizácia (vylepšenie obrázkov za pomoci ofarbenia)
- Vizualizácia pomocou ggplot2 (vytvorenie jedného profesionálneho obrazka na odprezentovanie)
- Záver

2 Zreplikujte nasledujúci obrázok (30b)



Dáta nájdete v knižnici `ggplot2` pod názvom `mpg` odozva je dojazd vozidla v míľach na dialsnici, voči objemu motora. Veľkosť bodu určuje dojazd vozidla v míľach v meste, ofarbenie určuje typ náhonu. V obrázku je potrebné dodatočne nastaviť:

- pozadie (theme)
- vypnúť legendu pre dojazd v meste (guides)
- prepísať názvy osí a legendy pre typ náhonu (labs)

3 (Bonus)

Nájdite príklad na *regression to mean* efekt, iný ako bol spomenutý na prednáške. Efekt podrobne vysvetlite aby bol zrozumiteľný aj neštatistikovi. V kreativite sa medze nekladú.