V souboru zelvyroznew.xlsx jsou údaje týkající se rozměrů želviček. Ve sloupci A je uvedena délka želvy (mm), ve sloupci B šířka (mm) a ve sloupci C výška (mm). V řádkách 1-48 jsou údaje týkající se želvích samců a v řádkách 49-96 jsou údaje týkající se samiček. Každý bude zpracovávat pouze řádky, které mu jsou určeny v souboru Zadání pro rok 2025 STM.

1. Nakreslete tři obrázky, v prvním budou dva boxploty (krabicové grafy), jeden pro délku samečků a jeden pro délku samiček. V druhém obrázku bude totéž pro šířku a ve třetím pro výšku. Je ve Vašich údajích o délce, respektive o šířce, respektive o výšce, nějaké odlehlé pozorování?
2. Předpokládejte, že Vaše data týkající se délky samečků, jsou výběrem ze základních souborů, který obsahuje všechny délky samečků tohoto typu. Obdobně Vaše data týkající se délky samiček, jsou výběrem ze základního souboru, který obsahuje všechny délky samiček tohoto typu. Udělejte odhad pro rozdíl průměrné délky samečků a samiček tohoto typu želv a sestrojte pro něj 95% interval spolehlivosti.
3. Spočtěte korelační matici pro délky, šířky a výšky, zvlášť pro samečky a zvlášť pro samičky. Které veličiny jsou nejvíc zkorelovány a které nejméně. Okomentujte. Domníváte se na základě Vašich dat, že je silnější korelace mezi délkou a šířkou želvy pro samce nebo samičky? Je korelace mezi délkou a šířkou, délkou a výškou, šířkou a výškou, silná nebo slabá?
4. Odhadněte průměrný rozdíl mezi délkou a šířkou samečka a zkonstruujte pro něj 95% interval spolehlivosti. Totéž udělejte pro samičky. Proč musíte použít párovou analýzu?
5. Předpokládejte, že o určitém samečkovi víme, že má délku 141mm a šířku 105mm? Jeho výšku neznáme. Jak bychom odhadli jeho výšku při použití lineární regrese se dvěma vysvětlujícími proměnnými? Zkonstruujte 95% zjednodušený predikční interval pro jeho výšku.
6. Předpokládejme o tomtéž samečkovi, že známe pouze jeho délku, která je 141 mm. S použitím jednoduché lineární regrese, kde vysvětlující proměnná je délka, odhadněte jeho výšku. Zkonstruujte opět 95% zjednodušený predikční interval pro jeho výšku. Porovnejte délky obou predikčních intervalů. Okomentujte.
7. Pokud bychom předpokládali, že zkoumaná želva je samička, přičemž údaje o délce a šířce jsou jako v bodu 5, tj. délka 141mm a šířka 105mm, byla by bodová predikce menší nebo větší než pro samce. Jaká by byla délka 95% predikčního intervalu? Kratší nebo delší? Vysvětlete.