EXAM PREPARATION 2

11. Anda tertarik untuk melihat apakah Total bill memiliki hubungan dengan besar uang tip yg dberikan. Berikut adalah hasil uji normalitas data anda:

```
normaltest(df['total_bill'])

NormaltestResult(statistic=45.11781912347332, pvalue=1.5951078766352608e-10)

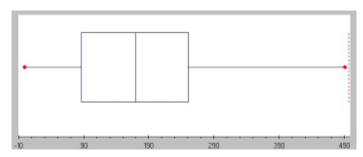
normaltest(df['tip'])

NormaltestResult(statistic=79.37862574074785, pvalue=5.796294322907102e-18)
```

Manakan code yg digunakan untuk melihat korelasi antara total bill dan tip

- a. df[['total_bill','tip']].corr()
- b. df[['total_bill','tip']].corr('spearman')
- c. df.corr()[['total_bill','tip']]
- d. df['total_bill'].corr(df['tip'])
- e. df['total_bill'].corr(df['tip'], method='pearson')
- 12. Uji Hipotesis yg paling cocok untuk pernyataan "Harga Mobil pabrikan jepang (Honda, Toyota, Daihatsu, suzuki, mitsubishi) di Indonesia sama besar"
 - a. T test Independent
 - b. Z Test
 - c. Wilcoxon
 - d. Kruskal wallis
 - e. Mann Whitney U
- 13. Pernyataan yg tepat mengenai "Two tail Test"
 - a. Ketika akan membandingkan hasil secara teori dan secara aktual
 - b. Ketika kita tidak mengetahui dan tidak dapat menentukan arah dari Hipotesis alternatif (lebih besar atau lebih kecil)
 - c. Two tail hanya digunakan untuk Uji parametrik
 - d. Two tail hanya digunakan untuk Uji Non Parametrik
 - e. Two tail tidak dapat digunakan untuk uji kruskal wallis Test
- 14. Tentukan jenis plot dan penggunaannya.
 - a. Pie plot digunakan untuk melihat proporsi data
 - b. Barplot digunakan untuk melakukan komparasi berdasarkan data kategorikal
 - c. Boxplot digunakan untuk melihat distribusi data dan outliers
 - d. A, B, C salah
 - e. A, B, C Benar
- 15. Pernyataan yg benar terkait distribusi normal adalah
 - a. Nilai mean dan median berbeda
 - b. Nilai mean lebih besar dari median
 - c. Nilai median lebih besar dari mean
 - d. Tidak ada nilai mean dan median
 - e. Persebaran data di dekat mean

- 16. Pernyataan yg salah terkait Uji statistik adalah
 - a. Wilcoxon dan kruskal wallis adalah contoh uji non parametrik
 - b. Uji non parametrik menggunakan median sebagai ukuran pusat data
 - c. Uji non parametrik digunakan ketika kita tidak memiliki cukup data
 - d. Anova dan T test adalah contoh uji parametrik
 - e. Ada 3 jenis uji statistik, parametrik, non parametrik, semi parametrik
- 17. Pernyataan yg benar mengenai Statistik deskriptif adalah
 - a. Jika distribusi data normal gunakan IQR untuk melihat variasi data
 - b. Standar deviasi digunakan untuk variasi data jika data tidak terdistribusi normal
 - c. Jika distribusi data anda Skewed gunakan median sebagai ukuran central tendency
 - d. Mean digunakan sebagai ukuran central tendency data kategorikal
 - e. Semua jawaban benar
- 18. Manakah di antara pernyataan berikut yg merupakan Hipotesis Alternatif
 - a. Tidak ada hubungan antara preferensi pria dan wanita terhadap kendaraan yg digunakan
 - b. Data pada kolom C mengikuti distribusi Normal
 - c. Terdapat perbedaan nilai TOEFL antara sebelum dan sesudah mengikuti pelatihan
 - d. Tes yg dikerjakan secara tertulis dan secara online membutuhkan waktu yg sama
 - e. Pria dan wanita memiliki rata-rata gaji yg sama
- 19. Jika kita ingin melihat hubungan antara 2 feature continous numerikal seperti Luas Lahan dan Harga rumah, grafik yg digunakan adalah
 - a. Pie Chart
 - b. Line Plot
 - c. Bar Plot
 - d. Scatter Plot
 - e. Box Plot
- 20. Perhatikan Boxplot berikut, jelaskan kondisinya



- a. Uniform
- b. Positive Skewed
- c. Negative Skewed
- d. Normal
- e. Tidak Skewed