

# ĐỀ THI CUỐI KỲ

**Môn học:** TKUD

**Hình thức:** Làm bài trên máy tính và nộp online

**Ngày thi:** 27/01/2024

**Thời gian:** 60 phút (9h45' – 10h45')

## I. LÝ THUYẾT (3.5 điểm)

- 1) Trình bày định nghĩa và phân loại của thống kê **(1,5 điểm)**
- 2) Trình bày khái niệm về các số đo thống kê mô tả số học sau đây: *phân vị, trung bình, phương sai, độ lệch chuẩn*. **(2 điểm)**

## II. BÀI TẬP (6 điểm)

Cho cơ sở dữ liệu điều tra về giá đất ở nông thôn trên địa bàn **huyện A** như bảng bên dưới.

<b>Y</b>	<b>H</b>	<b>G</b>	<b>T</b>	<b>R</b>
5,000,000	2.5	2,073.0	5,298.0	3.0
5,100,000	2.5	1,653.0	5,500.0	5.1
5,200,000	2.3	700.0	3,598.0	3.2
5,200,000	3.4	202.0	3,590.0	5.0
5,300,000	2.8	406.0	1,984.0	3.0
6,000,000	5.2	2,080.0	5,818.0	3.0
6,000,000	2.9	1,095.0	4,526.0	5.0
6,000,000	5.0	80.0	675.0	5.0
6,000,000	3.0	1,740.0	5,177.0	4.0
6,000,000	2.0	36.0	700.0	4.7
6,100,000	2.0	56.0	1,600.0	3.0
6,100,000	1.5	342.0	600.0	4.0

6,200,000	3.7	296.0	3,178.0	5.0
6,200,000	3.7	253.0	3,088.0	5.0
6,300,000	2.0	1,057.0	2,315.0	5.0
6,300,000	3.0	947.0	1,100.0	4.0
6,300,000	2.0	120.0	3,790.0	5.0
6,300,000	3.0	295.0	3,569.0	6.0
6,300,000	2.2	320.0	1,007.0	5.0
6,400,000	2.6	250.0	1,256.0	6.9

Dữ liệu trên phục vụ thiết lập mô hình hồi quy tuyến tính giá đất, gồm các thông tin và giả thiết như sau:

❖ **Thiết lập mô hình hàm giá đất tổng quát:  $Y = \alpha_0 + \alpha_1 H + \alpha_2 G + \alpha_3 T + \alpha_4 R$**

Trong đó:

+ Y: Giá đất (triệu đồng/m<sup>2</sup>)

+ H: Độ rộng hẻm (m)

+ G: Khoảng cách từ thửa đất đến đường giao thông đã được đặt tên hoặc đã được quy định trong BGĐ (m)

+ T: Khoảng cách từ thửa đất đến khu vực trung tâm xã hoặc chợ, trường học (m)

+ R: Độ rộng mặt tiền thửa đất (m)

+  $\alpha_0$  là hệ số tự do.

+  $\alpha_1$  đến  $\alpha_4$  là các hệ số các biến độc lập tương ứng theo thứ tự H, G, T, R

❖ **Giả thiết mối quan hệ giữa các biến:**

+ *Độ rộng hẻm (H)*: những thửa đất có độ rộng hẻm lớn, giao thông thuận lợi thì giá đất sẽ cao hơn, nên kỳ vọng là (+)

+ *Khoảng cách từ thửa đất đến đường giao thông đã được đặt tên hoặc đã được quy định trong BGĐ (G)*: những thửa đất xa đường giao thông hơn thì giá đất sẽ thấp hơn, nên kỳ vọng là (-)

+ *Khoảng cách từ thửa đất đến khu vực trung tâm xã hoặc chợ, trường học (T)*: những thửa đất xa khu vực trung tâm xã, chợ, trường học hơn thì giá đất sẽ thấp hơn, nên kỳ vọng là (-)

+ *Độ rộng mặt tiền thửa đất (R)*: những thửa đất có mặt tiền rộng hơn thì giá đất sẽ cao hơn, nên kỳ vọng là (+)

### **YÊU CẦU:**

1/ Nhập dữ liệu vào R, tạo dataframe lưu tên **stt. giadat** (trong đó, **stt** là số thứ tự của sinh viên, ghi 2 chữ số. Ví dụ: sinh viên có số thứ tự 01 sẽ tạo dataframe với tên là **giadat.01**). **(1 điểm)**

2/ Tính toán 4 số đo thống kê số học sau đây: *phân vị, trung bình, phương sai, và sai số chuẩn* cho biến giá đất (Y) trong cơ sở dữ liệu được cho ở trên. **(1 điểm)**

3/ Vẽ biểu đồ tần suất (histogram) thể hiện tần suất giá đất (Y) theo dạng cột với các thông số cơ bản như sau: biểu đồ cột nền màu vàng, đường viền màu đen, có thể hiện được đường xác suất màu đỏ. **(1 điểm)**

3/ Xây dựng mô hình hồi quy tuyến tính đa biến ( $Y = \alpha_0 + \alpha_1H + \alpha_2G + \alpha_3T + \alpha_4R$ ) xác định giá đất ở nông thôn của huyện A:

- Xác định được  $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4$  và mô hình hồi quy. **(1 điểm)**

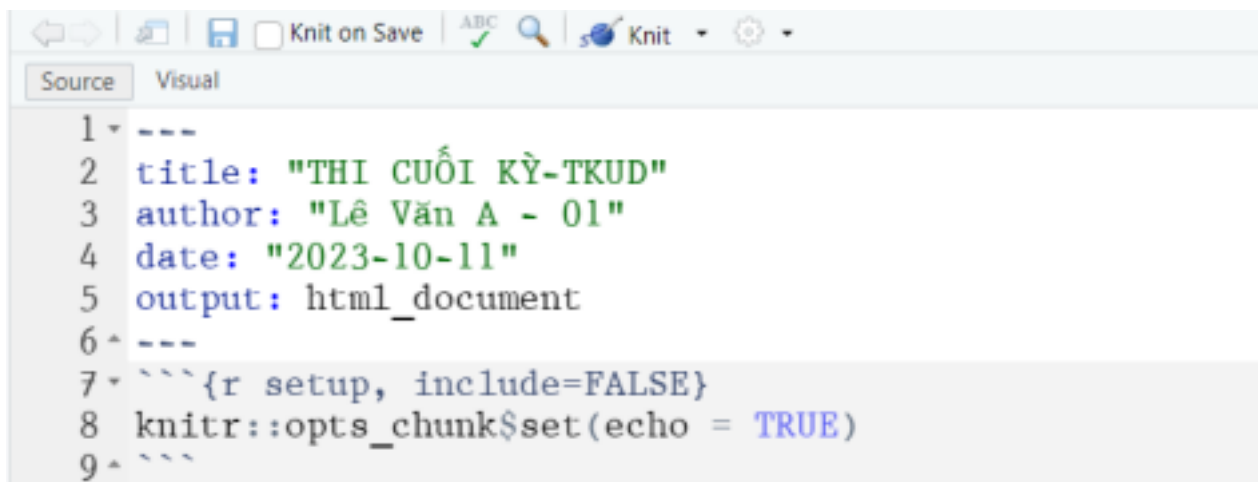
- Đánh giá mức ý nghĩa thống kê của các biến số hồi quy **(1 điểm)**

- Đánh giá và giải thích mức thích hợp (độ mạnh) của mô hình hồi quy sử dụng các hệ số  $R^2$  (R bình phương) và  $R^2$  điều chỉnh. **(1 điểm)**

### **III. LƯU Ý**

1/ Hình thức và trình bày **(0.5 điểm)**

2/ Format đầu file Rmarkdown với các thông tin như hình bên dưới, lưu ý mục author phải có đầy đủ **Họ và tên – Số thứ tự** (ví dụ: LÊ VĂN A – 01).





```
1 ---
2 title: "THI CUỐI KỲ-TKUD"
3 author: "Lê Văn A - 01"
4 date: "2023-10-11"
5 output: html_document
6 ---
7 ```{r setup, include=FALSE}
8 knitr::opts_chunk$set(echo = TRUE)
9 ```
```

2/ Tên file Rmarkdown được đặt dựa trên môn học, năm học và số thứ tự của sinh viên trong danh sách. Ví dụ: bạn sinh viên tên **Lê Thị Mọc** học lớp Kinh tế lượng căn bản

(TKUD) năm học 2023-2024 (DH23-24) và có số thứ tự là **01** thì ghi như sau: **TKUD\_DH23-24\_01** (xem hình file bên dưới). Sinh viên chỉ cần nộp file nén (dạng zip hoặc rar) trong đó có chứa 2 file nêu trên (ví dụ: **TKUD\_DH23-24\_01.html** và **TKUD\_DH23-24\_01.Rmd**) kết quả nộp sẽ như thế này). *Xem minh họa bên dưới.*

- File nén được nộp 1 trong 2 kiểu sau, kiểu nào cũng được như hình bên dưới:

 TKUD_DH23-24_01.rar	1/24/2024 12:55 AM	WinRAR archive	200 KB
 TKUD_DH23-24_01.zip	1/24/2024 12:54 AM	Compressed (zipped)...	213 KB

- Trong file nén phải có chứa 2 file như hình bên dưới:

Name	Type	Compressed size	Password ...	Size	Ratio	Date modified
 TKUD_DH23-24_01.html	Chrome HTML Document	211 KB	No	616 KB	66%	10/8/2023 3:00 PM
 TKUD_DH23-24_01.Rmd	RMD File	2 KB	No	4 KB	51%	10/8/2023 3:00 PM

3/ Sinh viên bắt đầu nhận đề thi trên google classroom vào lúc 9 giờ 45 phút. Làm bài và nộp bài (file nén) trước thời gian đến hạn (10 giờ 50 phút) ngày 12/01/2024. Lưu ý, sau 10 giờ 50 thì tự động không nhận bài, sinh viên sẽ không nộp được nữa, trường hợp này xem như không tham dự thi (bỏ thi) và bị 0 điểm, học lại.

Thủ Đức, ngày 27 tháng 01 năm 2024.



TS. Huỳnh Thanh Hiền