

**Học xong bài này, em sẽ:**

- ✓ Nêu được một số trải nghiệm của bản thân trong việc trích rút thông tin và tri thức hữu ích từ dữ liệu đã có thông qua sử dụng công cụ phân tích dữ liệu nâng cao của Excel.

### Nhiệm vụ 1. Trích xuất thông tin về ước tính kì hạn vay từ dữ liệu cho trước

**Yêu cầu:**

Khi quyết định vay một số tiền lớn và trả góp hằng tháng trong một thời gian dài, người vay cần lựa chọn phương án phù hợp với khả năng trả góp hằng tháng của mình. Kì hạn vay phụ thuộc vào số tiền trả góp hằng tháng và mặt bằng lãi suất lúc đi vay. Hãy xác định kì hạn vay dựa trên dữ liệu về mặt bằng lãi suất cho trước và số tiền trả góp hằng tháng.

**Hướng dẫn thực hiện:**

1. Xác định vấn đề

Xác định kì hạn vay dựa trên dữ liệu về mặt bằng lãi suất cho trước.

2. Thu thập dữ liệu

Các dữ liệu đầu vào cần có:

- Số tiền cần vay;
- Số tiền có thể trả góp hằng tháng;
- Lãi suất theo năm của một số ngân hàng khi đi vay.

3. Phân tích dữ liệu, trích xuất thông tin bằng công cụ phân tích dữ liệu nâng cao của Excel

Từ dữ liệu đầu vào có thể trích xuất thông tin về thời hạn vay phù hợp với khả năng trả góp bằng cách dùng hàm **PMT** kết hợp với công cụ phân tích **What-if Analysis** có sẵn trong Excel.

**PMT** (viết tắt của “payment”) là một hàm tài chính của Excel có thể dùng để tính khoản thanh toán định kì cho một khoản vay. Khi nhân **PMT** với số tháng sẽ là tổng số tiền phải trả trong suốt kì hạn của khoản vay. Excel gợi ý cú pháp hàm **PMT** như sau:

**PMT= (rate, nper, pv, [fv], [type])**

Trong đó, *rate* là lãi suất (không đổi trong suốt kì hạn), *nper* là kì hạn (số tháng), *p<sub>v</sub>* là giá trị hiện tại của khoản đầu tư, *[fv]* và *[type]* là các tham số không bắt buộc phải có.

Công cụ phân tích **What-If Analysis** cho phép người dùng thử các giá trị khác nhau cho các công thức. **What-If Analysis** bao gồm ba công cụ chính:

- **Scenario Manager**: cho phép tạo và quản lí các kịch bản khác nhau, mỗi kịch bản có các giá trị đầu vào riêng.

- **Data Table:** cho phép xem cách các giá trị đầu vào của một công thức thay đổi khi thay đổi một hoặc nhiều giá trị đầu vào.

- **Goal Seek:** cho phép tìm giá trị đầu vào cần thiết để đạt được một mục tiêu cụ thể.

Với yêu cầu của nhiệm vụ này, ta sử dụng công cụ **Data Table** để thấy số tiền phải trả mỗi tháng thay đổi như thế nào khi thay đổi lãi suất hoặc vừa thay đổi lãi suất vừa thay đổi kì hạn vay. Căn cứ vào dữ liệu sau khi phân tích, người vay có thể ước tính kì hạn vay theo khả năng trả góp hằng tháng.

#### 4. Ví dụ minh họa

a) Các dữ liệu đầu vào thu thập được:

- Số tiền cần vay: 500 triệu VNĐ;
- Số tiền có thể trả góp hằng tháng: khoảng 10 triệu VNĐ;
- Lãi suất theo năm khi đi vay dao động tùy ngân hàng, ví dụ có các mức là 6,5%; 7,0; 7,5%; 8,0%.

b) Các bước thao tác:

**Bước 1. Lập khối ô tính PMT.**

- Nhập số tiền vay vào ô **C4**.
- Nhập mức lãi suất vay vào ô **C5**.
- Nhập kì hạn vay vào ô **C6**.
- Nhập vào ô **C7** hàm **=PMT(C5/12, C6, -C4)**.

Hình 1. Lập khối ô tính PMT

	A	B	C
1			
2			Hàm PMT
3			
4		vay nợ	500.000.000 VNĐ
5		lãi suất (năm)	7%
6		kì hạn (tháng)	24
7		trả góp hằng tháng	22.386.289,55 VNĐ

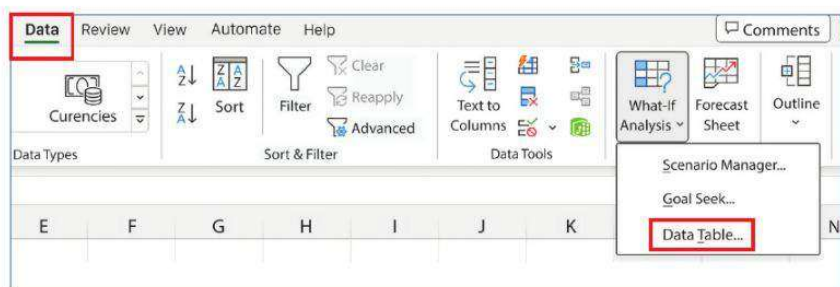
Vay tiền là đầu tư với giá trị âm nên để kết quả của hàm **PMT** ra một số dương ta dùng dấu trừ “-” trước tham số **pv** (**C4**). Kết quả nhận được như **Hình 1**.

**Bước 2. Phân tích dự báo What-If Analysis theo một biến lãi suất.**

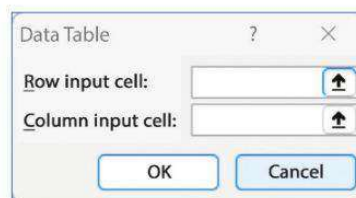
- Nhập dãy giá trị biến thiên của mức lãi suất trong một cột (hoặc hàng); ví dụ: khối ô **B13:B16**.

- Nhập vào ô **C12** hàm **=PMT(C5/12, C6, -C4)**.
- Đánh dấu chọn khối ô **B12:C16**.

- Chọn **Data**, chọn **What-If Analysis** trong nhóm lệnh **Forecast**. Trong bảng chọn thả xuống, chọn **Data Table** (**Hình 2**).



Hình 2. Chọn Data Table khi phân tích What-if Analysis



Hình 3. Hộp thoại Data Table



– Hộp thoại *Data Table* xuất hiện (Hình 3) nhập **\$C\$5** vào ô **Column input cell** và chọn **OK**. Kết quả nhận được như khối ô **B12:C16** trong Hình 4.

	A	B	C	D	E	F	G	H	I
11									
12		lãi suất (năm)	22,386,289.55 VND		22,386,289.55 VND	24	36	48	60
13		6.50%	22,273,125.72 VND		6.50%	22,273,125.72 VND	15,324,501.44 VND	11,857,476.46 VND	9,783,074.11 VND
14		7.00%	22,386,289.55 VND		7.00%	22,386,289.55 VND	15,438,548.43 VND	11,973,122.33 VND	9,900,599.27 VND
15		7.50%	22,499,796.33 VND		7.50%	22,499,796.33 VND	15,553,109.08 VND	12,089,450.97 VND	10,018,974.30 VND
16		8.00%	22,613,645.73 VND		8.00%	22,613,645.73 VND	15,668,182.73 VND	12,206,461.17 VND	10,138,197.14 VND

Hình 4. Phân tích *What-if Analysis* theo một biến (khối ô B12:C16) và theo hai biến (khối ô E12: I16)

**Bước 3.** Phân tích dự báo **What-If Analysis** theo hai biến lãi suất và kì hạn vay.

- Để dễ theo dõi, nhập dữ liệu cho khối ô **E13:E16** tương tự như **B13:B16**.
- Bổ sung thêm dãy giá trị biến thiên của kì hạn vay vào khối ô **F12: I12** để tạo thành khối ô hình chữ nhật.
- Nhập vào ô **E12** hàm **=PMT(C5/12, C6, -C4)**.
- Lặp lại các thao tác để xuất hiện hộp thoại *Data Table* như ở Bước 2.
- Nhập **\$C\$5** vào ô **Column input cell**.
- Nhập **\$C\$6** vào ô **Row input cell**.
- Chọn **OK**. Kết quả nhận được khối ô **E12:I16** trong Hình 4.

**Bước 4.** Trích xuất thông tin về ước tính kì hạn vay theo khả năng trả góp hằng tháng căn cứ vào kết quả nhận được. Giá trị trong khối ô **I13:I16** xấp xỉ 10 triệu VND. Kết luận: Thời hạn vay sẽ vào khoảng 60 tháng.

## Nhiệm vụ 2. Đưa ra dự báo dựa trên chuỗi thời gian

### Yêu cầu:

Xét ví dụ minh họa trình bày ở Bài 2 (trang 134). Để tìm ra mối quan hệ phụ thuộc của số lượng hành khách qua sân bay theo các chu kì thời gian, tổ dự án đã thu thập số liệu thống kê lượng hành khách hằng tháng trong quá khứ.

Giả sử tệp Excel chứa chuỗi thời gian gồm hai cột, cột A kiểu *Date* ghi lại chu kì thời gian (theo tháng) và cột B kiểu *Number* ghi lại số hành khách trong tháng đó (Hình 5).

Dựa trên chuỗi thời gian đó, sử dụng công cụ dự báo của Excel để:

- a) Xem kết quả dự báo và các tham số được thiết lập theo mặc định.
- b) Thay đổi một số tham số để hiểu và giải thích ý nghĩa của chúng trong kết quả dự báo nhận được.
- c) Rút ngắn chuỗi thời gian đầu vào, ví dụ bỏ bớt các tháng của năm 2013. Dùng chuỗi thời gian đã rút ngắn để dự báo, so sánh với số liệu thực tế đã có và cho nhận xét.

	A	B	
1	Date	Airport Passengers	Forecast
2	Jan-09	2,644,539	
3	Feb-09	2,359,800	
4	Mar-09	2,925,918	
5	Apr-09	3,024,973	
6	May-09	3,177,100	
7	Jun-09	3,419,595	
8	Jul-09	3,649,702	
9	Aug-09	3,650,668	
10	Sep-09	3,191,526	
11	Oct-09	3,249,428	
12	Nov-09	2,971,484	
13	Dec-09	3,074,209	
14	Jan-10	2,785,466	
15	Feb-10	2,515,361	
16	Mar-10	3,105,958	
17	Apr-10	3,139,059	
18	May-10	3,380,355	
19	Jun-10	3,612,886	
20	Jul-10	3,765,824	
21	Aug-10	3,771,842	
22	Sep-10	3,356,365	
23	Oct-10	3,490,100	
24	Nov-10	3,163,659	
25	Dec-10	3,167,124	
26	Jan-11	2,883,810	
27	Feb-11	2,610,667	
28	Mar-11	3,129,205	
29	Apr-11	3,200,527	
30	May-11	3,547,804	
31	Jun-11	3,766,323	
32	Jul-11	3,935,589	
33	Aug-11	3,917,884	
34	Sep-11	3,564,970	
35	Oct-11	3,602,455	
36	Nov-11	3,326,859	
37	Dec-11	3,441,693	
38	Jan-12	3,211,600	
39	Feb-12	2,998,119	
40	Mar-12	3,472,440	
41	Apr-12	3,563,007	
42	May-12	3,820,570	
43	Jun-12	4,107,195	
44	Jul-12	4,284,443	
45	Aug-12	4,356,216	
46	Sep-12	3,819,379	
47	Oct-12	3,844,987	
48	Nov-12	3,478,890	
49	Dec-12	3,443,039	
50	Jan-13	3,204,637	
51	Feb-13	2,966,477	
52	Mar-13	3,593,364	
53	Apr-13	3,604,104	
54	May-13	3,933,016	
55	Jun-13	4,146,797	
56	Jul-13	4,176,486	
57	Aug-13	4,347,059	
58	Sep-13	3,781,168	
59	Oct-13		

Hình 5. Một ví dụ chuỗi thời gian

### Hướng dẫn thực hiện:

a) Thực hiện các bước sau (cho yêu cầu a và c):

**Bước 1.** Chọn khối ô chứa dữ liệu chuỗi thời gian rồi chọn **Data**.

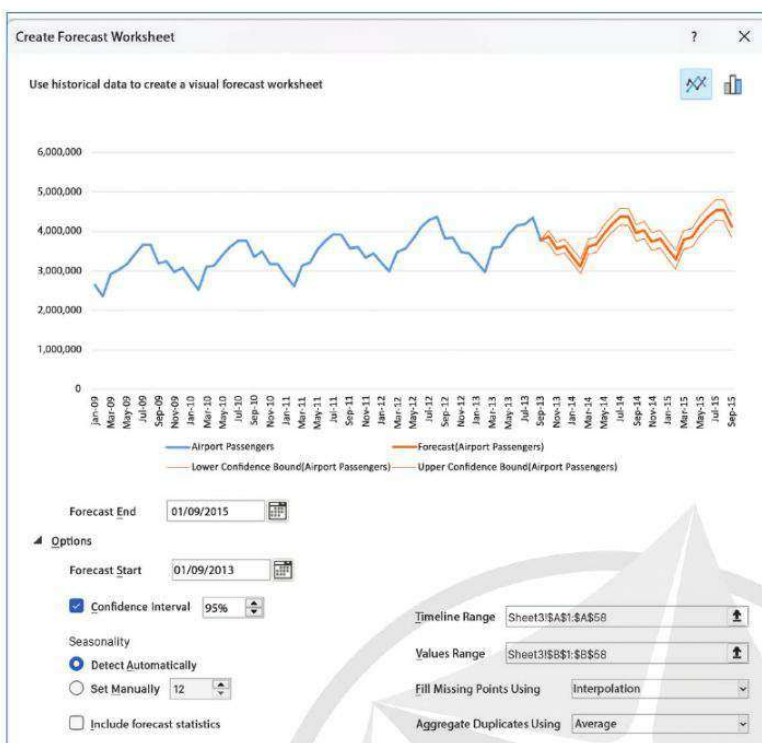
**Bước 2.** Chọn **Forecast Sheet** trong nhóm lệnh **Forecast**. Hộp thoại *Create Forecast Worksheet* xuất hiện.

**Bước 3.** Chọn **Create**, kết quả dự báo được tạo ra theo các thiết lập mặc định và lưu thành một trang mới.

**Bước 4.** Chọn **Options** để mở rộng hộp thoại *Create Forecast Worksheet* (Hình 6) và xem thiết lập mặc định cho các tham số: Forecast Start, Forecast End, Confidence Interval.

**Bước 5.** Kết quả thông tin được khai phá từ dữ liệu là dự báo cho một số tháng tiếp theo. Hình 6 là kết quả dự báo dựa trên tệp dữ liệu đã có dưới dạng biểu đồ đường. Đường màu xanh là biểu diễn dữ liệu hiện có (số lượng hành khách từ 01/01/2009 đến 01/09/2013), đường màu cam là biểu diễn dự đoán dữ liệu trong tương lai (số lượng hành khách từ 01/09/2013 đến 01/09/2015). Hình 7 là kết quả dự báo được trình bày trong khối ô C59:C82.





Hình 6. Kết quả dự báo dưới dạng biểu đồ đường dựa trên chuỗi thời gian trong Hình 5

59	Oct-13	3,858,196
60	Nov-13	3,562,680
61	Dec-13	3,633,798
62	Jan-14	3,366,457
63	Feb-14	3,110,903
64	Mar-14	3,614,670
65	Apr-14	3,666,432
66	May-14	3,960,805
67	Jun-14	4,182,886
68	Jul-14	4,367,447
69	Aug-14	4,363,455
70	Sep-14	3,954,015
71	Oct-14	4,031,044
72	Nov-14	3,735,527
73	Dec-14	3,806,646
74	Jan-15	3,539,305
75	Feb-15	3,283,750
76	Mar-15	3,787,518
77	Apr-15	3,839,280
78	May-15	4,133,652
79	Jun-15	4,355,733
80	Jul-15	4,540,295
81	Aug-15	4,536,303
82	Sep-15	4,126,863

Hình 7. Kết quả dự báo dựa trên chuỗi thời gian trong Hình 5

b) Lặp lại các thao tác như trên cho đến Bước 2. Tiếp theo, chọn **Options** để thay đổi một số tham số trước khi chọn **Create** xem kết quả và giải thích ý nghĩa.

Bỏ đánh dấu chọn **Confidence Interval** trước khi chọn **Create**, rút ra kết luận về tác dụng của lựa chọn này. Đánh dấu lựa chọn **ConfidenceInterval**, thay đổi tăng/giảm giá trị của **Confidence Interval**, chọn **Create**, cho biết tác động của sự thay đổi này tới đồ thị biểu diễn.



Nước ta xuất khẩu nhiều mặt hàng, trong đó có hải sản, rau quả là các mặt hàng có tính mùa vụ trong một năm. Hãy sưu tầm một chuỗi thời gian về xuất khẩu hải sản (hoặc rau quả) làm dữ liệu đầu vào và phân tích dự báo dựa trên chuỗi thời gian này để ước tính kim ngạch xuất khẩu trong năm tiếp theo.

### Gợi ý về nguồn dữ liệu:

– Tìm kiếm với cụm từ “số liệu xuất nhập khẩu các tháng năm 2023” để truy cập trang “[https://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2023/03/so-lieu-xuat-nhap-khau-cac-thang-nam-2023/...](https://www.gso.gov.vn/du-lieu-va-so-lieu-thong-ke/2023/03/so-lieu-xuat-nhap-khau-cac-thang-nam-2023/)”.

– Tìm mục “Tệp đính kèm”; nháy chọn, ví dụ “Trị giá và mặt hàng xuất khẩu chủ yếu sơ bộ các tháng của năm 2023 (.xls)” hoặc “Trị giá và mặt hàng nhập khẩu chủ yếu

sơ bộ các tháng năm 2023 (.xls)". Tập Excel chứa số liệu xuất khẩu (nhập khẩu) nhiều mặt hàng sẽ xuất hiện (Hình 8).

– Thao tác tương tự như trên, nhưng trong cụm từ tìm kiếm thay 2023 thành 2022 sẽ nhận được số liệu xuất nhập khẩu các tháng của năm 2022. Bằng cách lùi dần như vậy, có thể nhận được chuỗi thời gian dài hơn.

3	Tên hàng	Tháng 01		Tháng 02		Tháng 03		Tháng 4		Tháng 5	
		Lượng (Tấn)	Trị giá (1000 USD)	Lượng (Tấn)	Trị giá (1000 USD)	Lượng (Tấn)	Trị giá (1000 USD)	Lượng (Tấn)	Trị giá (1000 USD)	Lượng (Tấn)	Trị giá (1000 USD)
12	Hàng hải sản		457 212		607 610		764 766		741 756		807 754
13	Hàng rau quả		242 030		322 923		417 007		389 511		654 702

Hình 8. Minh họa số liệu xuất nhập khẩu các tháng năm 2023 của Việt Nam

## BÀI TÌM HIỂU THÊM

### PHẦN BỔ SUNG DATA MINING TRONG EXCEL

Excel có các phần bổ sung (Add-ins) giúp thực hiện phân tích dữ liệu nâng cao, bao gồm cả loại miễn phí và loại phải trả phí để có thể sử dụng. Data Mining là một Add-ins thực hiện khai phá dữ liệu có thể bổ sung miễn phí với Office 365. Các công cụ khai phá dữ liệu có sẵn sau khi bổ sung sẽ xuất hiện trong nhóm lệnh **Data Mining** thuộc dải lệnh **Data Mining** (Hình 9) gồm: phân loại (Classify), dự báo (Predict) và phát hiện luật kết hợp hay sự tương quan trong tập dữ liệu (Association).



Hình 9. Dải lệnh của phần bổ sung Data Mining trong Excel

Để thêm phần bổ sung **Data Mining** vào Excel, ta thực hiện theo các bước sau:

**Bước 1.** Trong cửa sổ làm việc của Excel, nhấp chuột chọn **Insert\Get Add-ins**.

**Bước 2.** Hộp thoại *Office Add-ins* xuất hiện, tìm và chọn phần bổ sung muốn có.

**Bước 3.** Đọc qua các thông tin cần biết. Sau đó, chọn **Add** (với các phiên bản phải trả phí, chọn **Try** để dùng thử, chọn **Buy** để thanh toán tiền mua).

**Bước 4.** Cần đăng nhập (sign in) bằng tài khoản Office 365 và chọn **Continue** để có thể sử dụng.

**Bước 5.** Mở dải lệnh **Data Mining** mới xuất hiện để xem kết quả.