التمرين الأول: عموميات

1: افتح برنامج بايثون واكتب ما يلي، ثم اكتشف دور **print**:

```
print("Hello World")
```

2: اكتب في برنامج بايثون ما يلي ثم اكتشف دور علامة **=**:

```
text = "Data analysis is important"
print(text)
```

3: اكتب في برنامج بايثون ما يلي ثم اكتشف دور **len**:

```
text = "Hello World"
print(len(text))
```

4: اكتب في برنامج بايثون ما يلي ثم اكتشف دور **lower**:

```
text = "HELLO PYTHON"
print(text.lower())
```

5: اكتب في برنامج بايثون ما يلي ثم اكتشف دور **split**:

```
text = "Data analysis is important"
words = text.split()
print(words)
```

6: اكتب في برنامج بايثون ما يلي ثم اكتشف دور **التعليمتين الثانية و الثالثة**:

```
text = "Data analysis is important"
words = text.split()
print(len(words))
```

7: اكتب في برنامج بايثون ما يلي ثم اكتشف دور **replace**:

```
text = "This is a text"
text = text.replace("is", "")
print(text)
```

التمرين الثاني: استخدام المكتبات

قبل استعمال أي مكتبة لأول مرة نقوم بتثبيتها في موجه الأوامر (Invite de commandes/ Command Prompt)، مثلا لتثبيت مكتبة ntlk نكتب : pip install nltk ثم قبل كل استعمال نستورد المكتبة في بايثون:

```
import nltk
nltk.download('punkt')
```

في الإصدارات الحديثة لبايثون نضيف أيضا: nltk.download('punkt_tab')

اكتب في بايثون ما يلي واستنتج دور التعليمات:

:1

```
import nltk

nltk.download('punkt')
text = "Data analysis is important"
words = nltk.word_tokenize(text)
print(words)
```

:2

```
import nltk
nltk.download('punkt')
text = "Hello World"
tokens = nltk.word_tokenize(text)
print(len(tokens))
```

:3

```
import nltk
import string
text = "HELLO PYTHON"
lower_text = text.lower() #
print(lower_text)
```

التمرين الثالث: الفرق في النتيجة بين استخدام المكتبات و دونها

1: اكتب في بايثون و لاحظ الفرق:

```
text = "Python, is easy."
print(text.split())

----------

import nltk
print(nltk.word_tokenize(text))
```

:2

```
text  = "مرحبا، كيف حالك؟ أنا أتعلم البرمجة."
words = text.split()
print(words)
```

```
print(len (words))
```

```
import nltk
nltk.download('punkt')
text = "مرحبا، كيف حالك؟ أنا أتعلم البرمجة."
words = nltk.word_tokenize(text)
print(words)
print(len(words))
```

التمرين الرابع: تحليل النصوص

اكتب في بايثون ما يلي ولاحظ النتيجة ثم اكتشف دور التعليمات:

```
text = """أصبحت معالجة النصوص من المجالات المهمة في علوم الحاسوب،
حيث تُستخدم لتحليل البيانات النصية واستخراج المعلومات المفيدة منها.
كما تساعد البرمجة بلغة بايثون على تنفيذ عمليات متعددة، مثل تقسيم النص إلى كلمات،
وحساب عدد الكلمات، وحذف الكلمات غير المهمة، واستخراج الكلمات الأكثر تكرارًا."""
```

```
print(text)
import nltk
nltk.download('punkt')
words = nltk.word_tokenize(text)
print(words)
print(len(words))
clean_words = [word for word in words if word.isalpha()]
print(clean_words)

from collections import Counter
freq = Counter(clean_words)
print(freq)
print(freq.most_common(3))
```