**تكليف علم بيانات محاضرة 1**

* **ما هي أهمية تحليل البيانات وعلم البيانات بشكل عام بالنسبة للشركات التجارية والمؤسسات الحكومية والقطاعات المختلفة في بلادنا؟ دلل بالأمثلة الواقعية الملموسة بدون الاستعانة بالذكاء الصناعي؟**

**\*أهمية تحليل البيانات وعلم البيانات:**

**-تحسين عملية صنع القرار:** يساعد تحليل البيانات المؤسسات على اتخاذ قرارات مبنية على الحقائق والأدلة بدلاً من الافتراضات أو الحدس.

**-خفض التكاليف وزيادة الكفاءة:** يساهم تحليل البيانات في تحديد المجالات التي يمكن فيها تقليل النفقات وتحسين العمليات التشغيلية، مما يؤدي إلى زيادة الإنتاجية وتقليل الهدر.

**-فهم العملاء واكتشاف فرص جديدة:** يمكن للشركات فهم تفضيلات العملاء وسلوكياتهم بشكل أفضل، مما يمكنها من تقديم خدمات ومنتجات مخصصة واكتشاف أسواق جديدة.

**-تحسين إدارة المخاطر:** يساعد تحليل البيانات في التعرف على التهديدات المحتملة والتنبؤ بها، مما يمكن المؤسسات من تطوير استراتيجيات لتقليل تأثير تلك التهديدات.

**-تعزيز الميزة التنافسية:** من خلال الرؤى المستخلصة من البيانات، يمكن للمؤسسات الاستجابة بفعالية لتغيرات السوق واكتساب ميزة تنافسية.

**-التنبؤ بالاتجاهات المستقبلية:** يستخدم علم البيانات تقنيات متقدمة مثل التعلم الآلي للتنبؤ بالاتجاهات المستقبلية، مما يساعد الشركات على الاستعداد للتحديات المحتملة.

**\*أمثلة واقعية ملموسة:**

**1-في الشركات التجارية:**

**-أمازون:** تستخدم أمازون تحليل البيانات بشكل مكثف لتحسين إدارة المخزون وتخصيص تجربة التسوق لكل عميل. فمن خلال تحليل بيانات التصفح والشراء، تقدم الشركة توصيات مخصصة للمنتجات، مما يزيد من فرص البيع.

**-سبوتيفاي:** تستخدم سبوتيفاي تحليل البيانات لإنشاء قوائم تشغيل مخصصة بناءً على عادات الاستماع لكل مستخدم، مما يحسن تجربة المستخدم ويزيد من ولائه.

**-نتفليكس:** تعتمد نتفليكس على تحليل البيانات لتحسين تجربة المشاهدة وتطوير المحتوى، من خلال فهم تفضيلات المستخدمين وأنماط المشاهدة.

**-ستاربكس:** تستخدم ستاربكس تحليل البيانات لتحسين تجربة العملاء وزيادة المبيعات، على سبيل المثال من خلال تحليل بيانات الشراء لتحديد العروض الترويجية المناسبة.

**-أوبر:** تستخدم أوبر تحليل البيانات لتحسين الكفاءة اللوجستية وخدمة العملاء، مثل تحسين مسارات السائقين وتحديد أوقات الذروة.

**2-في المؤسسات الحكومية والقطاعات المختلفة:**

**-القطاع الحكومي:** يمكن للمؤسسات الحكومية استخدام تحليل البيانات لتحسين الخدمات العامة بناءً على بيانات المواطنين، مثل تحديد أنماط المستفيدين وتفضيلاتهم لتطوير الخدمات وتحسين الاستراتيجيات.

**-قطاع النقل:** يمكن لهيئات النقل استخدام تحليل البيانات الضخمة حول كثافة حركة المرور وسرعتها وظروف الطقس للتنبؤ بخطر الحوادث وتنبيه السائقين، مما يساهم في نظام نقل ذكي.

**-القطاع المالي والمصرفي:** تساعد البيانات الضخمة في إدارة المخاطر واكتشاف الاحتيال، وتحسين خدمات البنوك وقرارات الاستثمار. على سبيل المثال، تستخدم البنوك تحليل البيانات الضخمة لتحليل المخاطر المالية والكشف عن الأنشطة الاحتيالية.

**-الرعاية الصحية:** يتعامل الأطباء والمستشفيات مع كميات هائلة من البيانات من سجلات المرضى، ويساعد تحليل البيانات في تحسين الرعاية الصحية، مثل تحليل الأشعة الطبية واكتشاف الأدوية.  
 **-الأمن السيبراني:** يعتمد الأمن السيبراني على علم البيانات لجمع معلومات قابلة للتنفيذ للمساعدة في تأمين الأنظمة والشبكات والبيانات بشكل أفضل.

* **مع تطور مصادر البيانات بشكل كبير وتزايد يوميا كم هو معدل زيادة البيانات بشكل سنويا(ابحث عن ذلك من مصادر علمية موثوقة)؟**تقرير صادر عن الأمم المتحدة، من المتوقع أن تزيد كمية البيانات عالميًا بأكثر من خمسة أضعاف، حيث سترتفع من 33 زيتابايت في عام 2018 إلى 175 زيتابايت بحلول عام 2025.

كما تشير إحصائيات أخرى من Edge Delta إلى أنه بحلول عام 2025، من المتوقع أن تصل البيانات الضخمة إلى 181 زيتابايت. ويُتوقع أن ينمو حجم سوق البيانات من 220.2 مليار دولار في عام 2023 إلى 401.2 مليار دولار بحلول عام 2028**.**

* **اذكر اهم خمسة تقنيات ومنهجيات وبرمجيات رائدة في مجال تحليل البيانات؟**

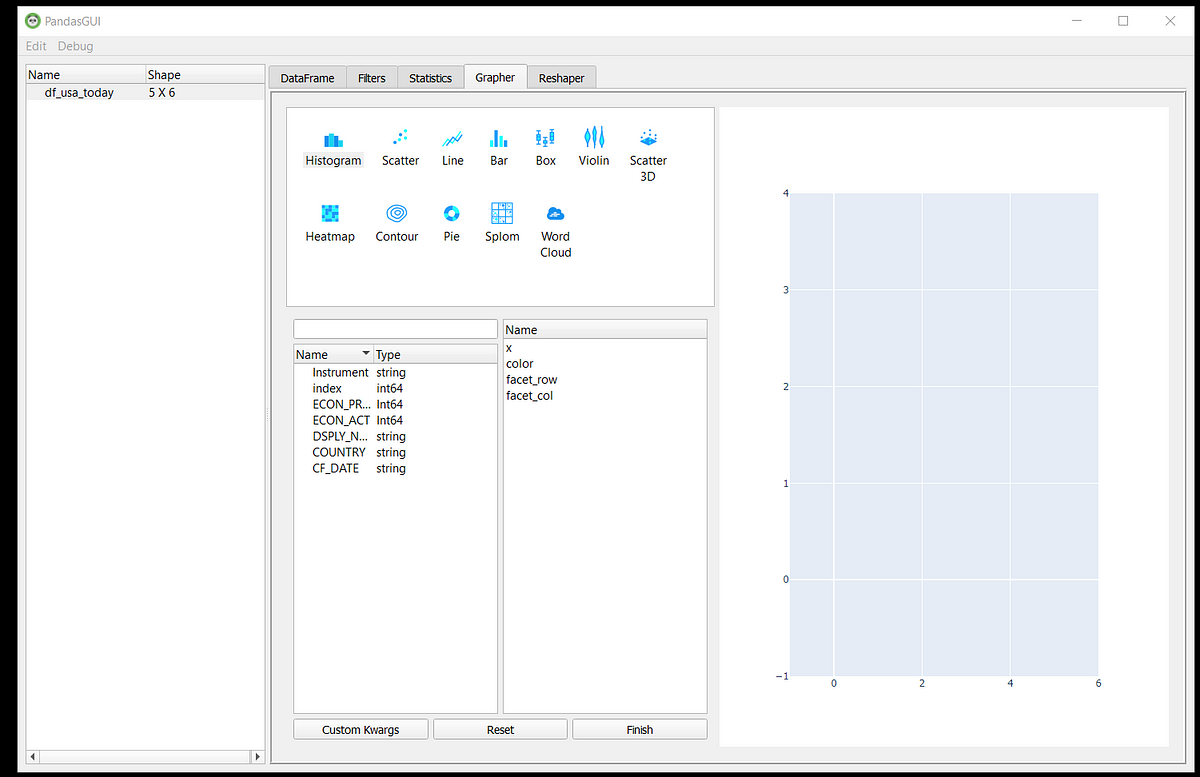
**1- Python**

تعتبر بايثون (Python) لغة البرمجة الأكثر شعبية في مجال علم البيانات، وذلك بفضل سهولة استخدامها، ومكتباتها الغنية، ومجتمعها الكبير والنشط. توفر بايثون بيئة متكاملة للتعامل مع البيانات، بدءًا من جمعها وتنظيفها، مرورًا بتحليلها ونمذجتها، وصولًا إلى تصور النتائج.



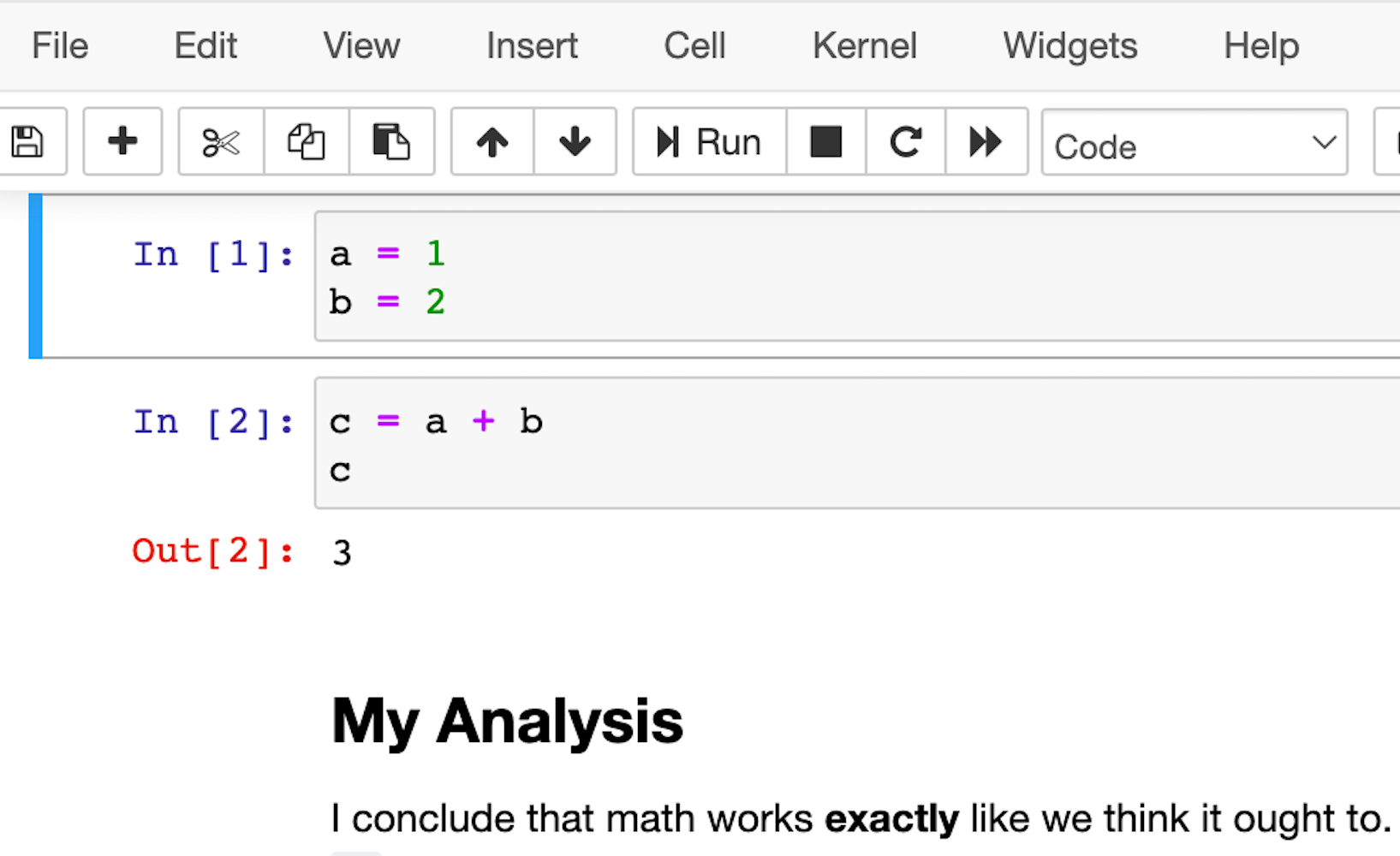
**2- Pandas**

تعد مكتبة Pandas من المكتبات الأساسية في بايثون لتحليل البيانات ومعالجتها. توفر هياكل بيانات قوية ومرنة مثل DataFrames، والتي تسهل التعامل مع البيانات الجدولية، وإجراء عمليات مثل التصفية، والتجميع، والدمج، وتحويل البيانات بكفاءة عالية.



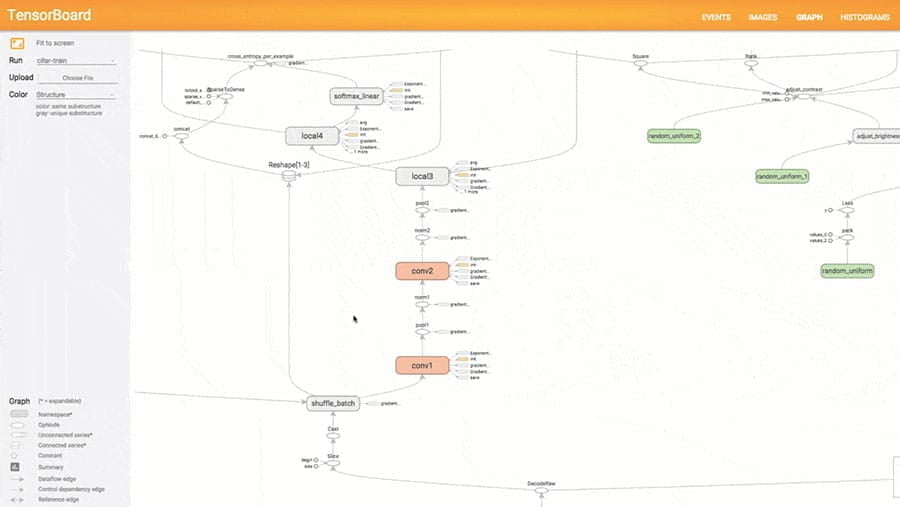
**3- NumPy**

هي مكتبة أساسية أخرى في بايثون، توفر دعمًا للمصفوفات الكبيرة متعددة الأبعاد والكائنات المصفوفية، بالإضافة إلى مجموعة واسعة من الدوال الرياضية للعمليات على هذه المصفوفات. تعتبر NumPy حجر الزاوية للعديد من مكتبات علم البيانات الأخرى، بما في ذلك Pandas وScikit-learn.



**4-TensorFlow**

هو إطار عمل مفتوح المصدر لتطوير ونشر نماذج التعلم الآلي والتعلم العميق. تم تطويره بواسطة Google، وهو يوفر أدوات قوية لبناء وتدريب الشبكات العصبية المعقدة، ويدعم الحوسبة الموزعة على وحدات المعالجة المركزية (CPUs) ووحدات معالجة الرسوميات(GPUs) وحتى وحدات المعالجة المخصصة(TPUs).



**5-Scikit-Learn**

هي مكتبة تعلم آلة مفتوحة المصدر مبنية على NumPy وSciPy. توفر مجموعة واسعة من خوارزميات التعلم الآلي للتصنيف، والانحدار، والتجميع، وتقليل الأبعاد، واختيار النماذج. تتميز بسهولة الاستخدام والتوثيق الشامل، مما يجعلها خيارًا ممتازًا للمبتدئين والخبراء على حد سواء.

