

## دایکومنت پروژه پایانی آزمایشگاه پایگاه داده

بهزاد فولادی ۴۰۰۷۳۱۲۹

### توضیح دیتابیس

هدف این پروژه، تحلیل شاخص کل بورس ایران در بازه زمانی سال ۲۰۰۸ تا ۲۰۲۳ با استفاده از پایگاه داده و کوئری‌های SQL است. این دیتابیس برای بررسی روند کلی بازار، تغییرات روزانه، نوسانات ماهانه و شناسایی نقاط شکست (Breakout) در شاخص بورس طراحی شده است. پایگاه داده شامل یک جدول اصلی بوده و تمرکز آن بر تحلیل داده‌ها به صورت آماری و زمانی هستش.

### کاربردها:

- تحلیل روند بلندمدت شاخص بورس
- محاسبه سود و زیان روزانه
- بررسی نوسان‌پذیری بازار
- شناسایی نقاط تغییر روند (صعودی و نزولی)

### ساختار دیتابیس

ساختار ساده و تحلیلی است و شامل یک موجودیت اصلی می‌باشد:

### جدول Overall\_Index :

نام ستون	توضیح
DTYYYYMMDD	تاریخ معاملات
OPEN	مقدار شاخص در شروع روز
HIGH	بیشترین مقدار شاخص در روز
LOW	کمترین مقدار شاخص در روز
CLOSE	مقدار شاخص در پایان روز
VOL	حجم معاملات
OPENINT	اطلاعات بهره باز (Open Interest)

از آن جا که هدف پروژه تحلیل داده‌هاست، رابطه‌ای بین چند جدول تعریف نشده و تمام تحلیل‌ها روی همین جدول انجام می‌شود.

## دستورالعمل نصب و راه‌اندازی پروژه

نرم‌افزارهای مورد نیاز :

- Microsoft SQL Server
- SQL Server Management Studio (SSMS)
- Microsoft Excel

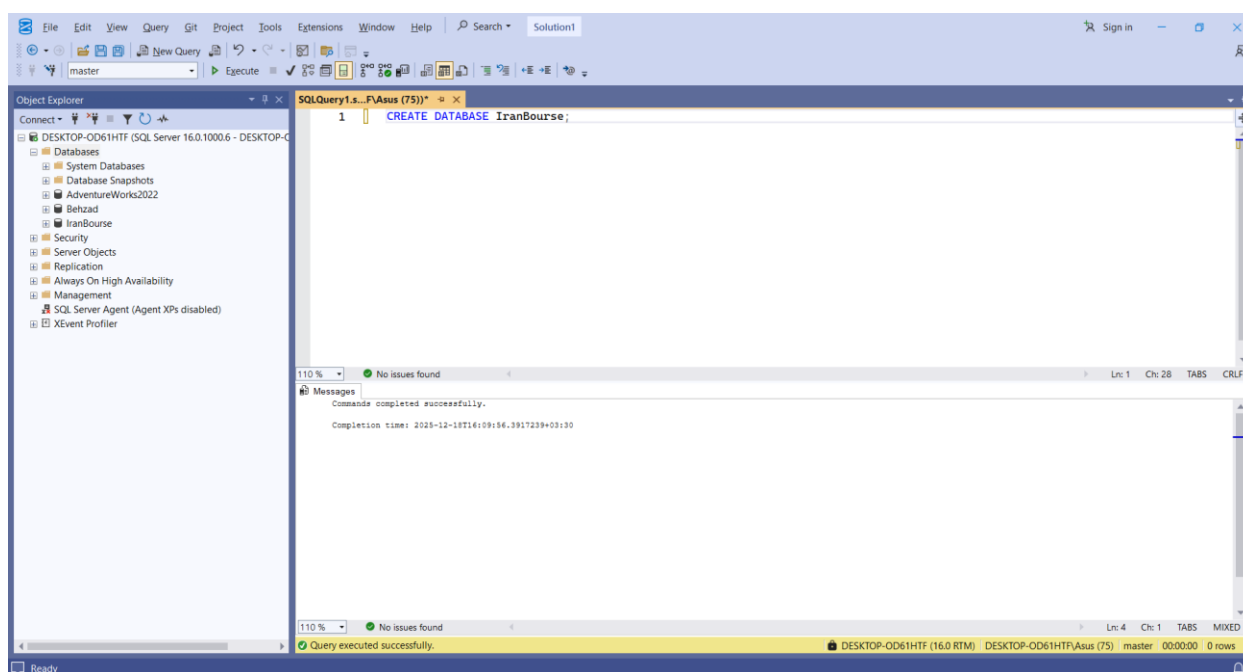
## مراحل راه‌اندازی

۱ . دانلود دیتاست از سایت Kaggle با این

لینک (<https://www.kaggle.com/datasets/nimapourmoradi/iran-bourse-overal-index/data>) البته در خود فولدر docs هم فایل excel موجود هست

۲ . باز کردن فایل Excel و تبدیل آن به فرمت CSV

۳ . ایجاد یک دیتابیس جدید در SQL Server



۴. استفاده از گزینه **Import Flat File** در SSMS برای وارد کردن فایل CSV (به عکس ها و نکات توجه شود).

The screenshot shows the 'Specify Input File' step of the 'Import Flat File' wizard. The left sidebar contains a navigation pane with the following items: Introduction, Specify Input File (selected), Preview Data, Modify Columns, Summary, and Results. The main area is titled 'Specify Input File' and contains the following fields:

- Location of file to be imported:** A text box containing 'D:\Projects\LABS\Data\Final\_Project\Overall\_Index.csv' and a 'Browse...' button.
- New table name:** A text box containing 'Overall\_Index'.
- Table schema:** A dropdown menu with 'dbo' selected.

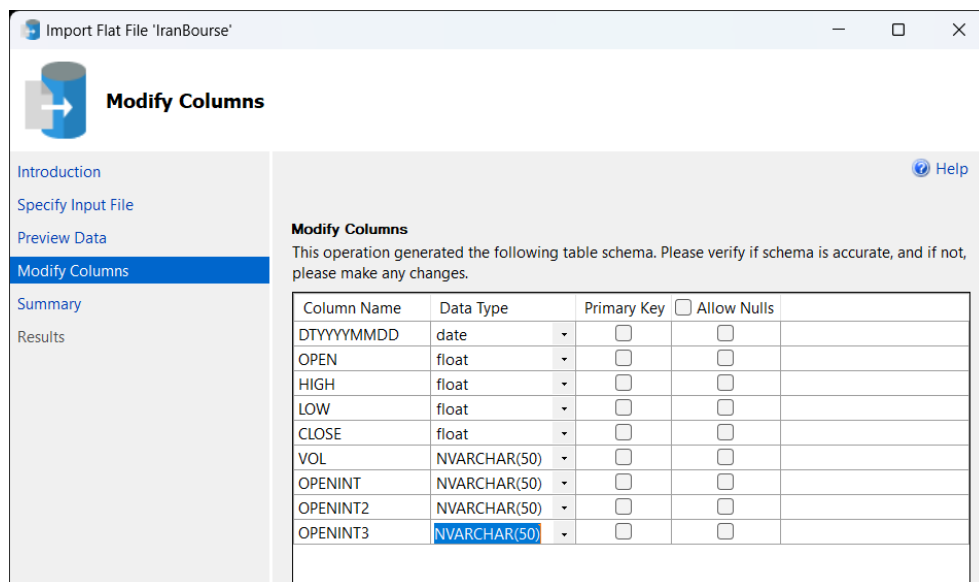
The screenshot shows the 'Preview Data' step of the 'Import Flat File' wizard. The left sidebar contains a navigation pane with the following items: Introduction, Specify Input File, Preview Data (selected), Modify Columns, Summary, and Results. The main area is titled 'Preview Data' and contains a table preview of the first 50 rows of the data.

**Preview Data**  
This operation analyzed the input file structure to generate the preview below for up to the first 50 rows.

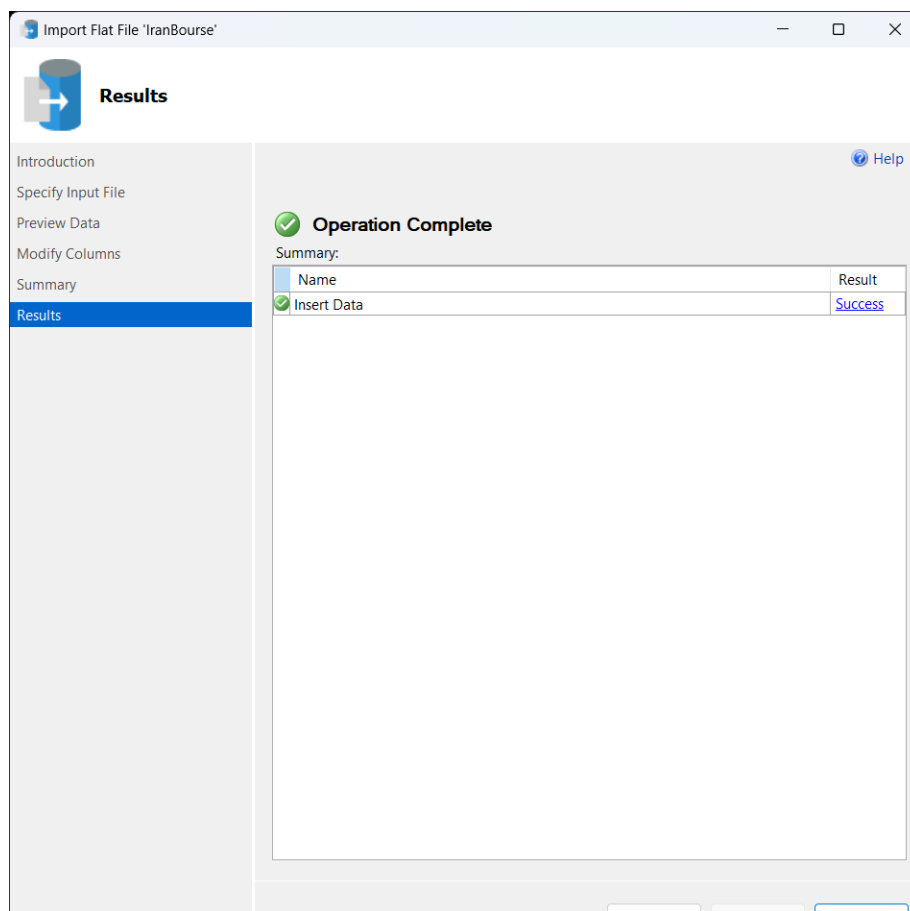
DTYYYYMMDD	OPEN	HIGH	LOW	CLOSE	
20081206	9248.8	9248.8	9167.8	9178.3	8
20081207	9178.3	9178.3	9130.3	9130.5	1
20081208	9102.7	9103.4	9079.6	9089.2	1
20081210	9071.6	9071.6	9023.7	9023.7	1
20081213	8973.3	8973	8973	8973.3	3
20081214	8963.9	8964.5	8907	8907	3
20081215	8901.8	8903.6	8890.7	8890.7	1
20081216	8890.8	8890.8	8866.3	8869	1
20081220	8861.5	8861.9	8854.2	8857.1	9
20081221	8857.3	8860.8	8851.6	8851.6	2
20081222	8840.1	8840.6	8832.2	8832.2	1
20081223	8829	8829.5	8803.4	8803.4	1
20081224	8796.4	8802.3	8792.5	8795.7	1
20081227	8784.9	8795.6	8784.9	8786.7	3
20081228	8769.7	8769.7	8757.8	8757.8	2
20081229	8750.1	8751.9	8726.2	8726.9	3
20081230	8717.2	8717.2	8689.8	8695	2

**Column names changed due to invalid characters, duplication, etc. Column names can be edited in Modify Columns page.**

☒ Use Rich Data Type Detection - may provide a closer type fit. However, cells with anomalous values may be dropped.



در قسمت Modify Columns باید دیتاتایپ ها را مانند تصویر وارد کنید تا در روند اجرای کوئری های به مشکل بر نخورید (علت استفاده از NVARCHAR در OPENINT ها این است اطلاعات اضافی هستند و کاربرد تحلیلی ندارند)



۵. اجرای کوئری‌های SQL ارائه‌شده برای تحلیل داده‌ها

تمام کوئری‌ها در فولدر sql به‌صورت مستقیم در SSMS قابل اجرا هستند و فقط باید در خود دیتابیس انتخابی Execute شوند.

## مثال داده‌ها

چند رکورد نمونه از جدول Overall\_Index :

DTYYYYMMDD	OPEN	HIGH	LOW	CLOSE	VOL
20081206	9248.8	9248.8	9167.8	9178.3	8539624
20081207	9180.1	9250.4	9175.6	9223.9	9123456

این داده‌ها نمایانگر اطلاعات روزانه شاخص بورس بوده و مبنای تمام تحلیل‌های پروژه هستند.

## اجرای کوئری‌های ساده

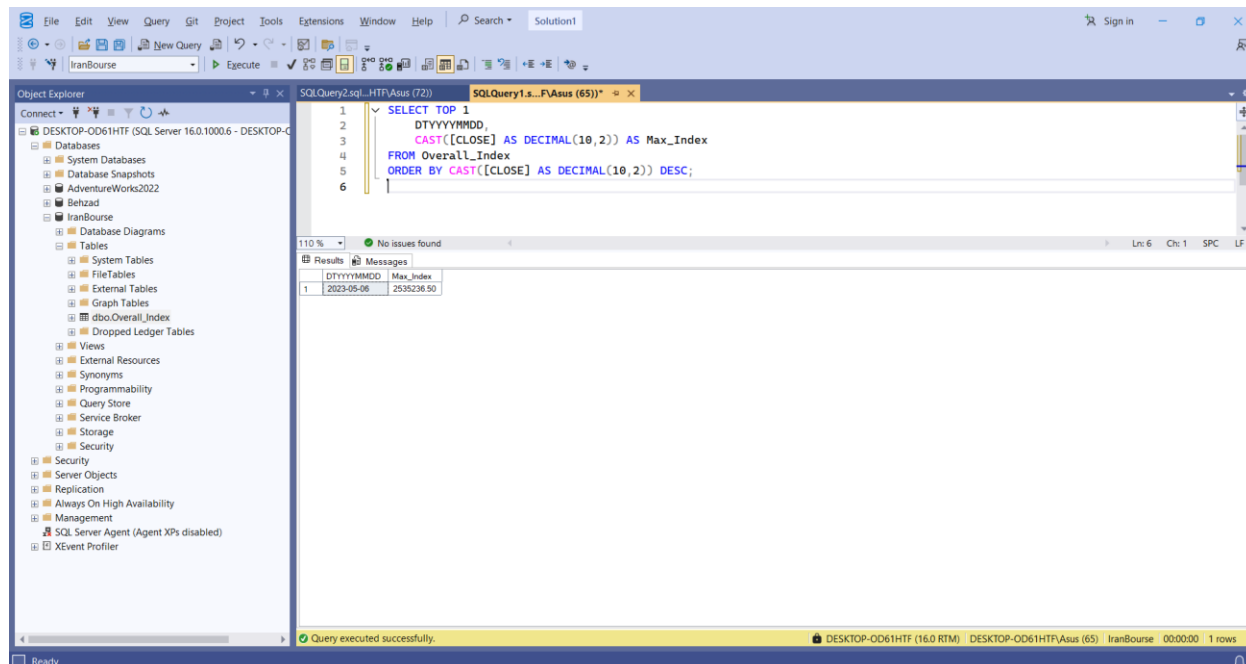
نمایش کامل رکوردهای جدول شاخص کل:

The screenshot shows the SQL Server Enterprise Manager interface. The query window displays the following SQL query:

```
1 SELECT *
2 FROM Overall_Index;
3
```

The Results pane shows the execution results in a table with the following columns: DTYYYYMMDD, OPEN, HIGH, LOW, CLOSE, VOL, OPENINT, OPENINT2, OPENINT3. The table contains 28 rows of data, starting from 2008-12-06 and ending on 2009-01-18. The status bar at the bottom indicates 'Query executed successfully.' and '3,587 rows'.

پیدا کردن بالاترین مقدار تاریخی شاخص:



## نتیجه گیری

در این پروژه، داده‌های تاریخی شاخص کل بورس ایران از سال ۲۰۰۸ تا ۲۰۲۳ در یک پایگاه داده مبتنی بر SQL Server ذخیره و تحلیل شدند. با استفاده از کوئری‌های ساده و تحلیلی، روند کلی بازار، تغییرات روزانه، نوسانات ماهانه و نقاط مهم تغییر روند شناسایی گردید. نتایج به‌دست‌آمده نشان می‌دهد که استفاده از پایگاه داده و زبان SQL ابزار قدرتمندی برای تحلیل داده‌های مالی و تصمیم‌گیری مبتنی بر داده فراهم می‌کند. این پروژه می‌تواند مبنایی مناسب برای تحلیل‌های پیشرفته‌تر و توسعه سیستم‌های تحلیلی در حوزه بازار سرمایه باشد.