**Данни от: Април 7 2023**

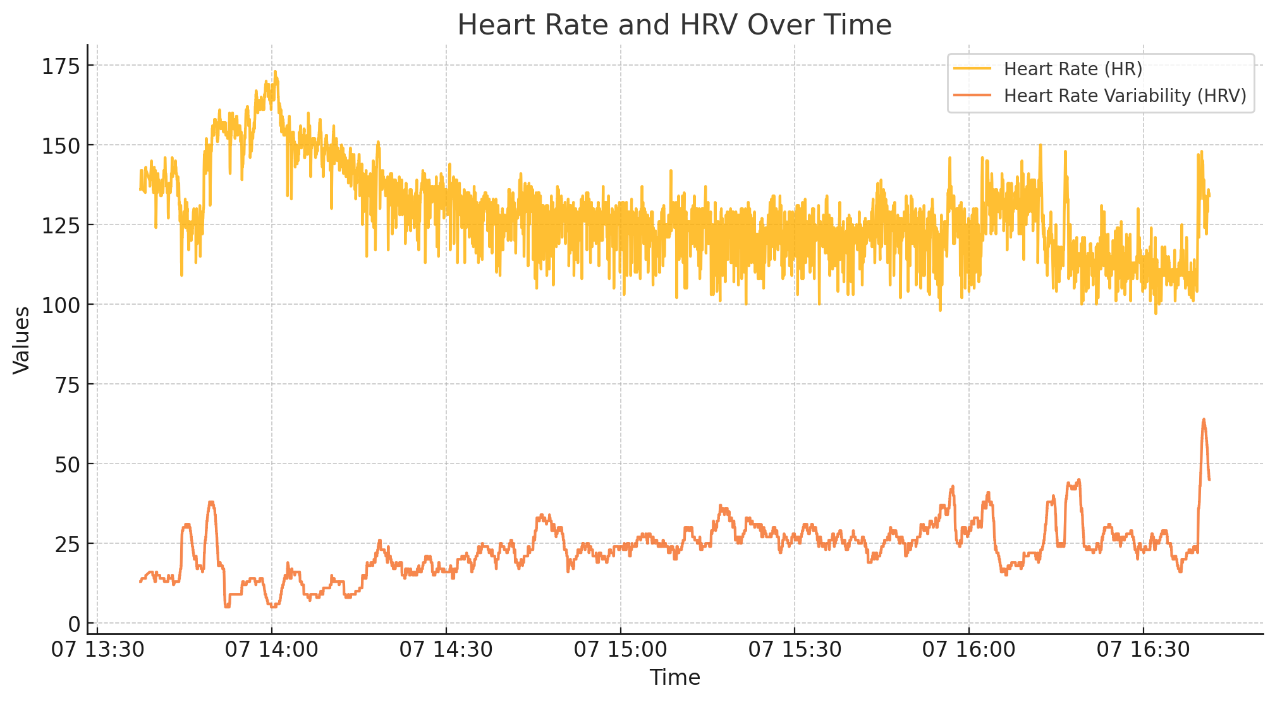
**Методология**

**1. Преглед и предварителна обработка на данните**  
Данните са заредени от предоставен файл и съдържат три основни колони: времеви марки (Time), сърдечен ритъм (HR) и вариабилност на сърдечната честота (HRV). Първоначалните стъпки включват:

* Конвертиране на времевите марки във формат datetime за правилна времева обработка.
* Интерполация на липсващи стойности (ако има такива) с цел създаване на непрекъснати времеви редове.
* Статистическо обобщение на данните, включващо изчисляване на средни стойности, стандартни отклонения и откриване на екстремни стойности.

**2. Визуализация на данните**

* **Линейна визуализация**: Изготвени са времеви графики, показващи тенденциите на HR и HRV в динамика.



Промени на вариабилността на сърдечната честота (HRV) и пулс (HR) във времето.

**Основни статистически показатели**  
Средната стойност на HR е 128.36, а на HRV – 23.33. Стандартните отклонения показват умерена променливост и в двете метрики (13.38 за HR и 8.10 за HRV).

**Обобщение на констатациите:**

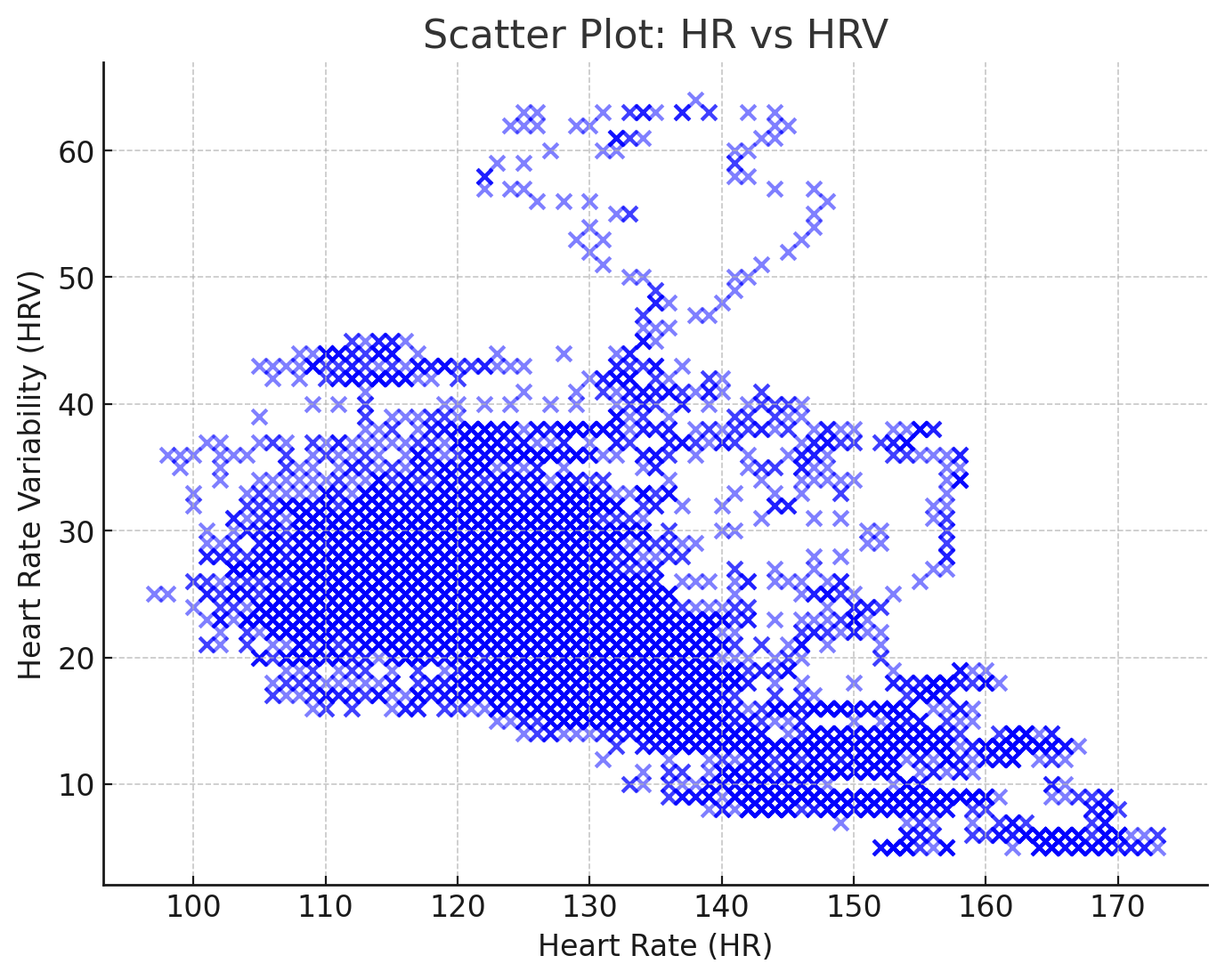
Средна сърдечна честота: 128.36

Средна вариабилност на сърдечния ритъм: 23.33

Стандартно отклонение на HR: 13.38

Стандартно отклонение на HRV: 8.10

**Скатер плот**: Използван за идентифициране на евентуална връзка между HR и HRV.

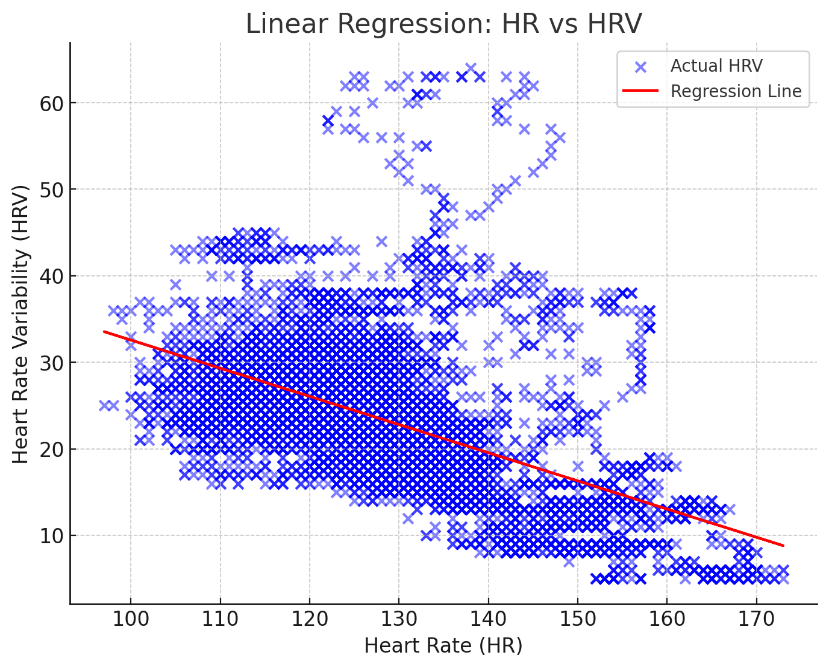


Точкова диаграма: Илюстрира връзката между HR и HRV.

**3. Корелационен анализ**  
Изчислени са две форми на корелация:

* **Пиърсън корелация (-0.538)**: Измерване на линейната зависимост между HR и HRV - налице е умерено отрицателна линейна връзка между HR и HRV, което показва, че с увеличаването на HR, стойностите на HRV намаляват.
* **Спиърман корелация (-0.556)**: За откриване на нелинейни, но монотонни зависимости. Потвърждава наличието на нелинейна, но стабилна зависимост.

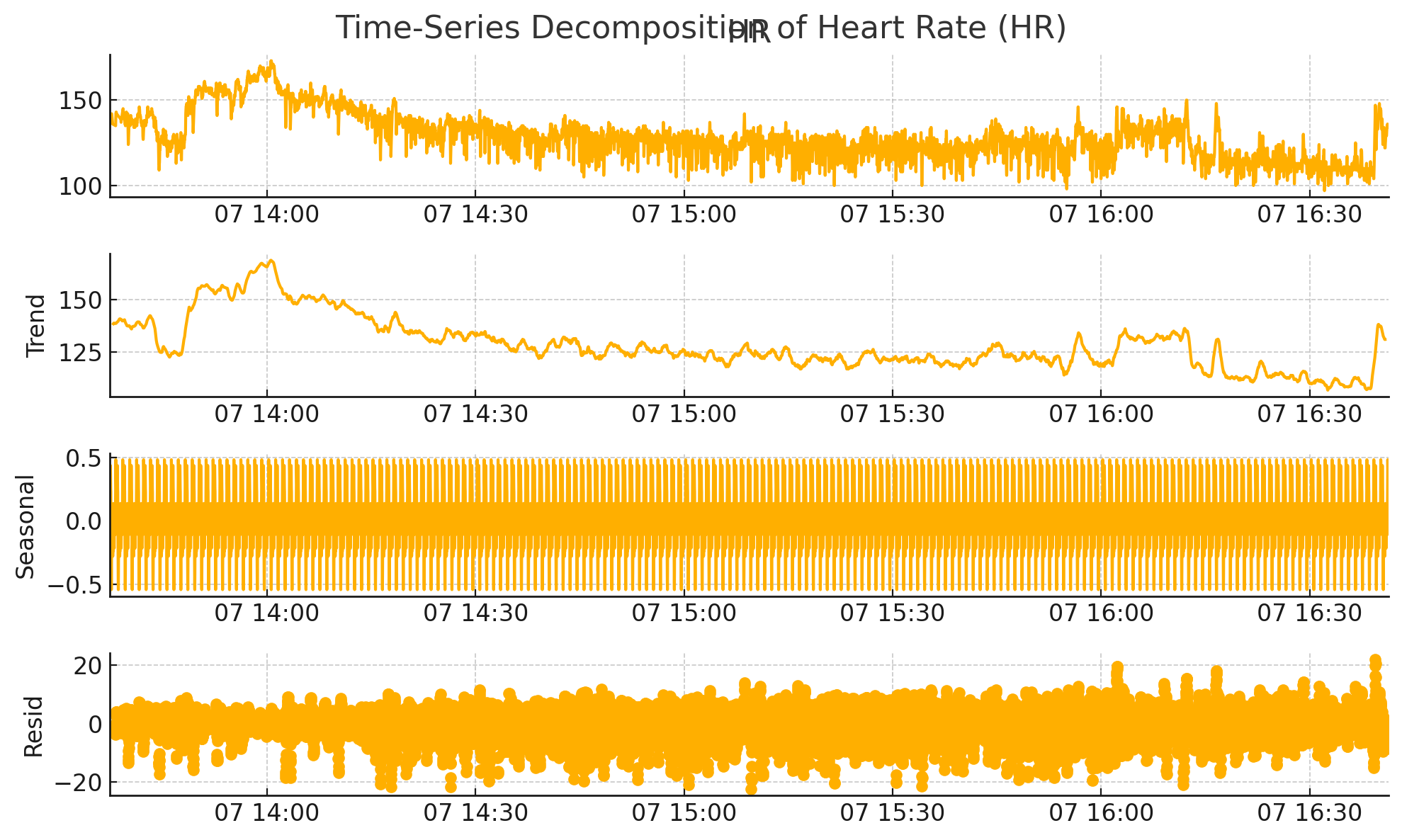
**4. Регресионен анализ**  
Линейната регресия е използвана за моделиране на връзката между HR (независима променлива) и HRV (зависима променлива). Моделът предоставя коефициенти на наклон и пресечна точка, които позволяват интерпретация на връзката.



**Линейна регресия**  
Коефициентът на наклон (-0.326) показва, че всяко увеличаване на HR с единица води до средно намаление на HRV с 0.326 единици. Регресионната линия обяснява част от променливостта, но не изцяло, което подсказва необходимост от допълнителни фактори.

**5. Времеви анализ**  
Данните са анализирани чрез сезонно разлагане (seasonal decomposition), за да се идентифицират тенденции, сезонни колебания и остатъчни компоненти.

Сезонното разлагане разкри ясна тенденция на спадове и пикове, които може да се свържат с периодични влияния (например физическа активност или стрес).



**6. Машинно обучение**  
За прогнозиране на HRV на база HR е използван Random Forest Regressor. Разделянето на данните е направено на тренировъчна (80%) и тестова (20%) извадка, като са изчислени метрики за точност (MSE и R²).

Random Forest моделът показва умерена прогностична точност:

* **MSE (48.77)**: Подчертава грешките в прогнозите.
* **R² (0.29)**: Означава, че моделът обяснява около 29% от променливостта на HRV.

