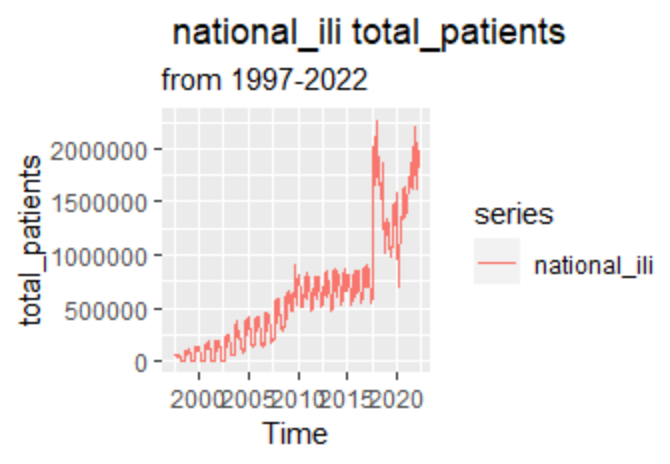
**חיזוי בעזרת תוכנת R- חלק 2:**

נשים לב לנתונים שהורדנו.

בקובץ נפרסים נתונים של חולים בטווחי גילאים שונים, שנים, שבועות ,המשקל שלהם ,ומספר חולים.

ניעזר בתצוגה גרפית כדי להסיק על התנהגות הנתונים :

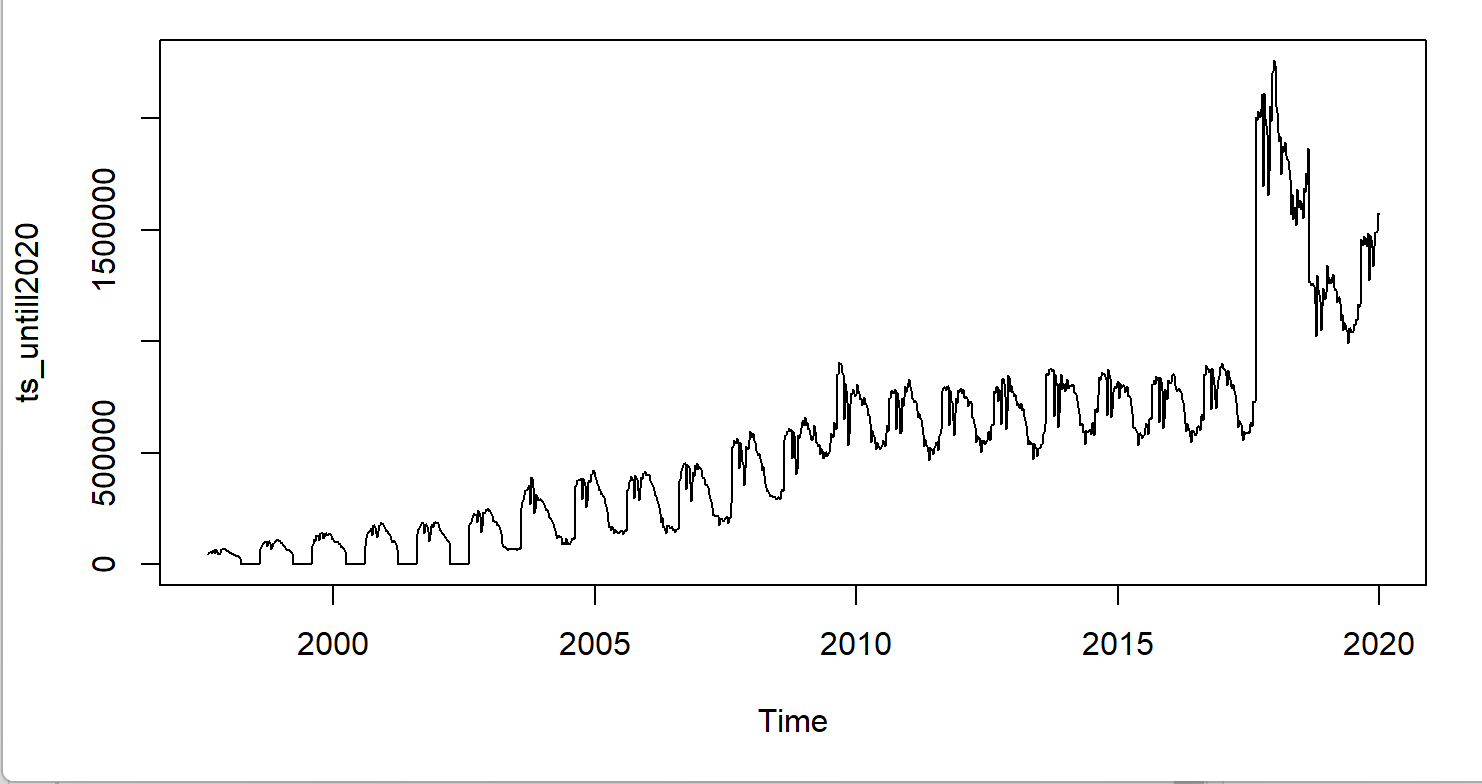


* נשים לב, יש שנים שאין להן מגמתיות קבועה, בהשוואה לקובץ האקסל ניתן לראות כי קיימים מקטעים עם נתונים חסרים בעמודות השונות (0).
* כמו כן, שנת 2022 מסתיימת בשבוע 18, סוף הדאטה שלנו.
* בשנת 2018 ישנה עלייה חדה בגרף- נסיק על עלייה חדה וחריגה במספר חולי השפעת /התפרצות המחלה בצורת דרסטית.
* נעבוד על סמך מגמתיות בכל מקרה, כיוון שמלבד 2018 העלייה מתבצעת בצורה די קבועה/שיטתית.

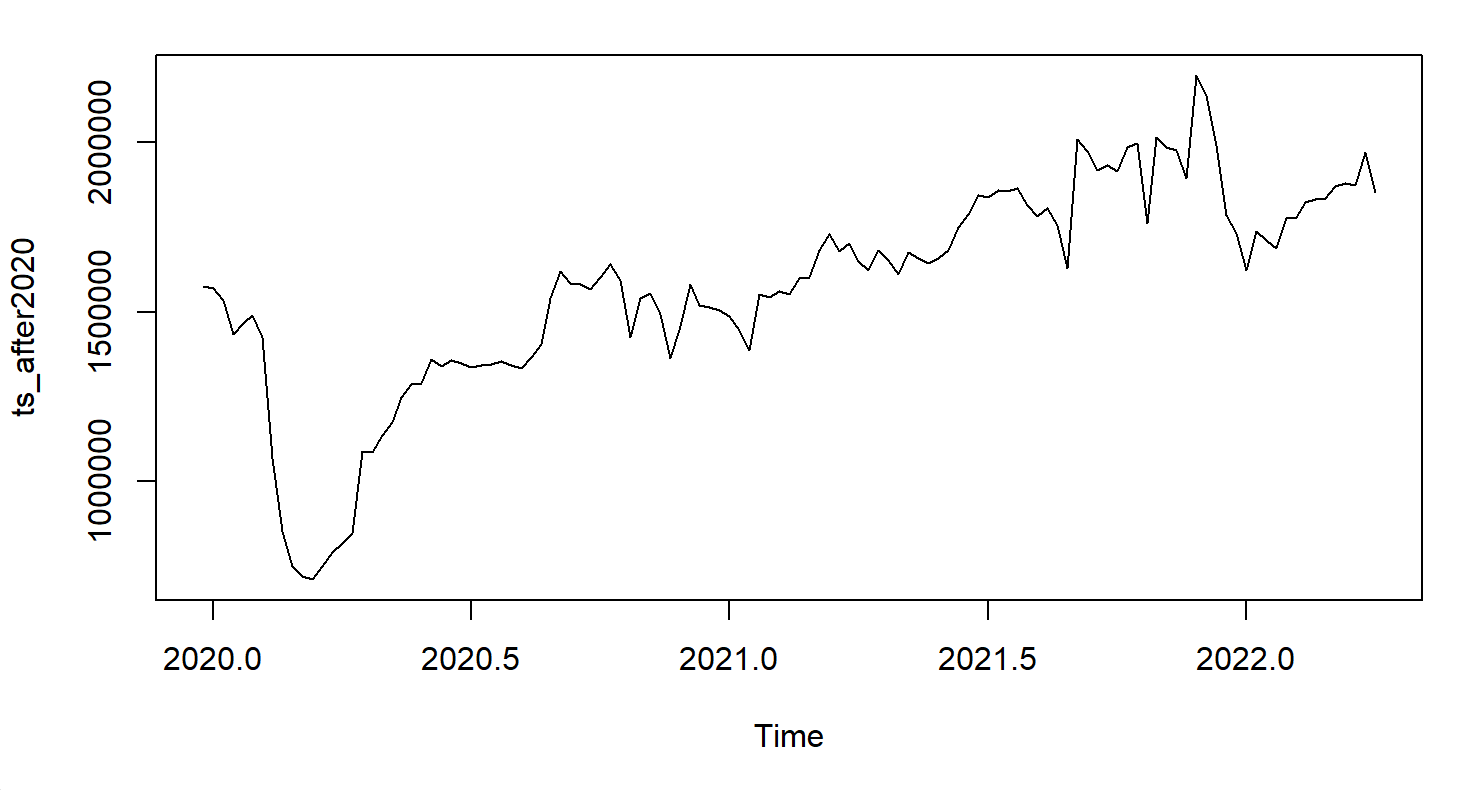
**סעיף C:**

נדרש: פיצול הסדרה לשני חלקים, פיצלנו לאחרי 2019 (לאחר סוף 2018) ולפני 2019 .

מאגר הנתונים לטובת אימון(train ):



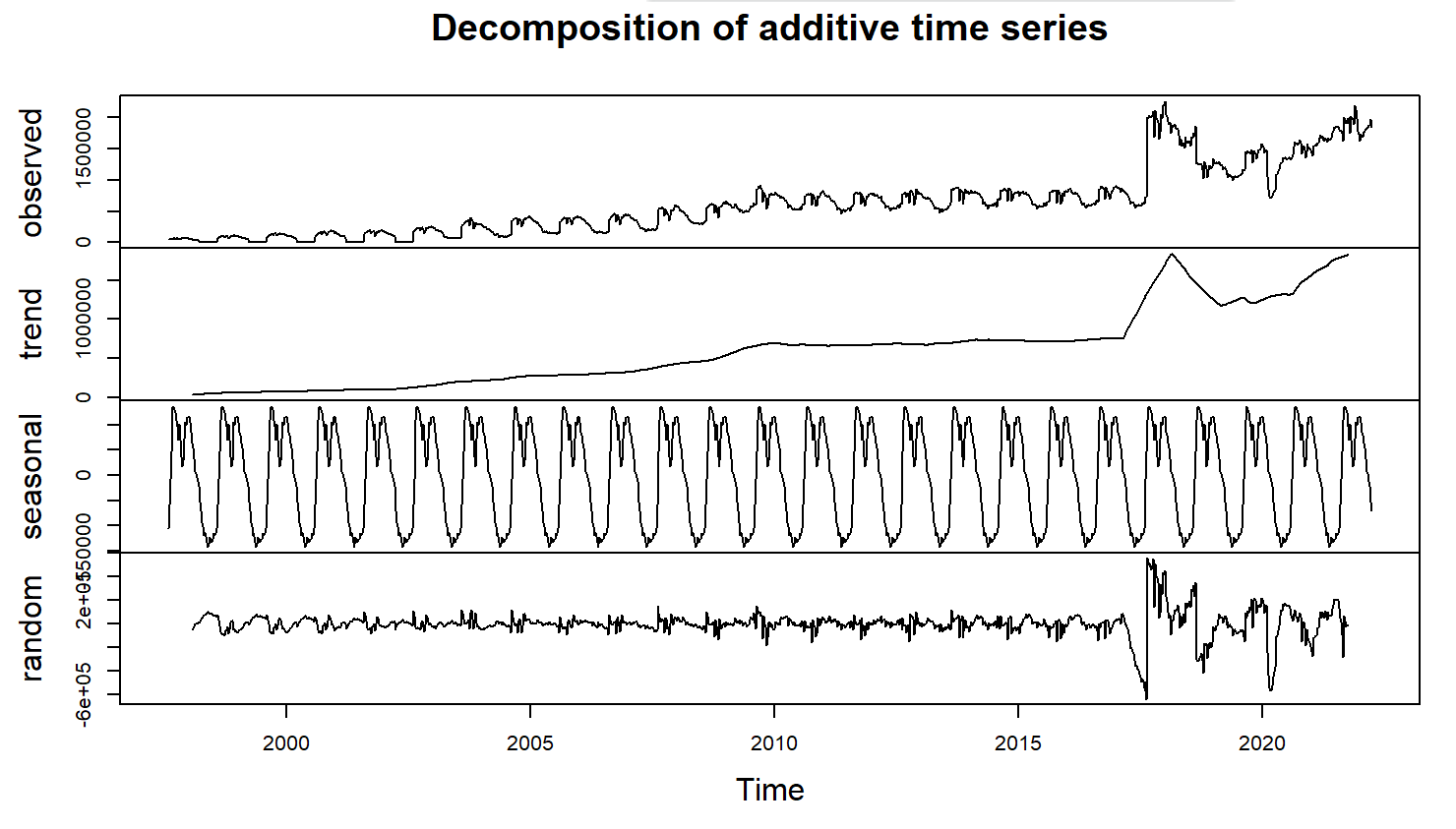
מאגר הנתונים לטובת המבחן: (test)



* **מכיוון שמצוין כי זו שיטה נהוגה במדעי הנתונים, נעזרנו בקורס הרלוונטי כדי לחלק את הבדיקה ל: אימון:80% ומבחן 20%**
* **האי דיוק שקיבלנו בהשוואה לממוצע נע הוא: 347633.7**

**סעיף D:**

נבחן מודלים שנלמדו בקורס, ניעזר בבחינת העונתיות בעזרת הולט ווינטרס **ודה קומפוזיציה** כדי "לפרק" לחלקים: נבחון קו מגמה,עונתיות, רעש וכמובן מציאות.

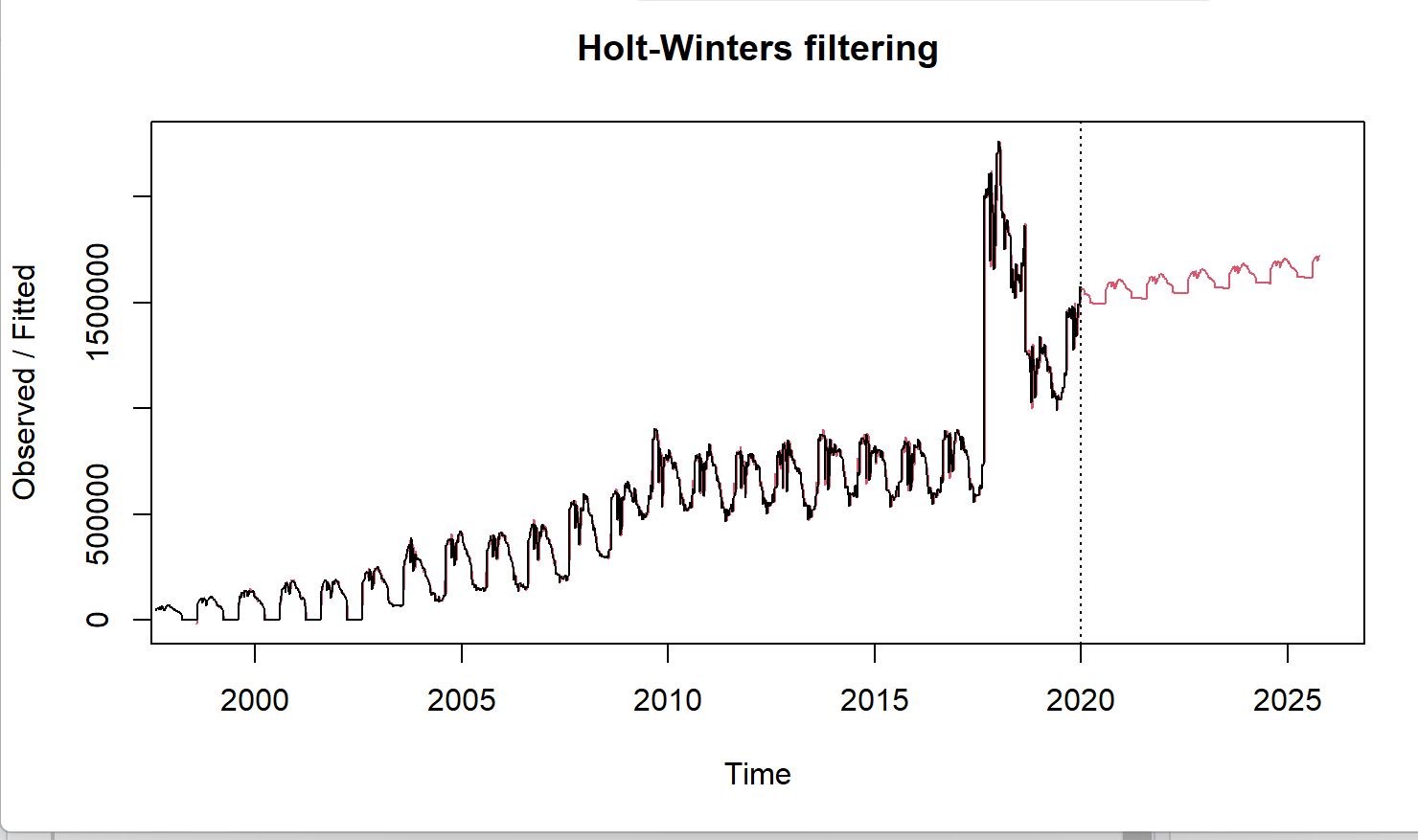


ניתן להבחין בעונתיות ברורה, מגמה עולה וברעש גבוה בשנת 2018.

* כעת נשתמש במודל **הולט ווינטרס** על הדאטה שלנו ונבדוק את ערכי: a- קבוע הישר מתבסס על הערכים.

B – השיפוע שמושפע מהתקופות, "נמעיט" בערכו בעקבות השינוי ב2018. וg-מתבסס על העונתיות,נשים לב שהdefult של התוכנה היא בחירתם כערכים אופטימליים.

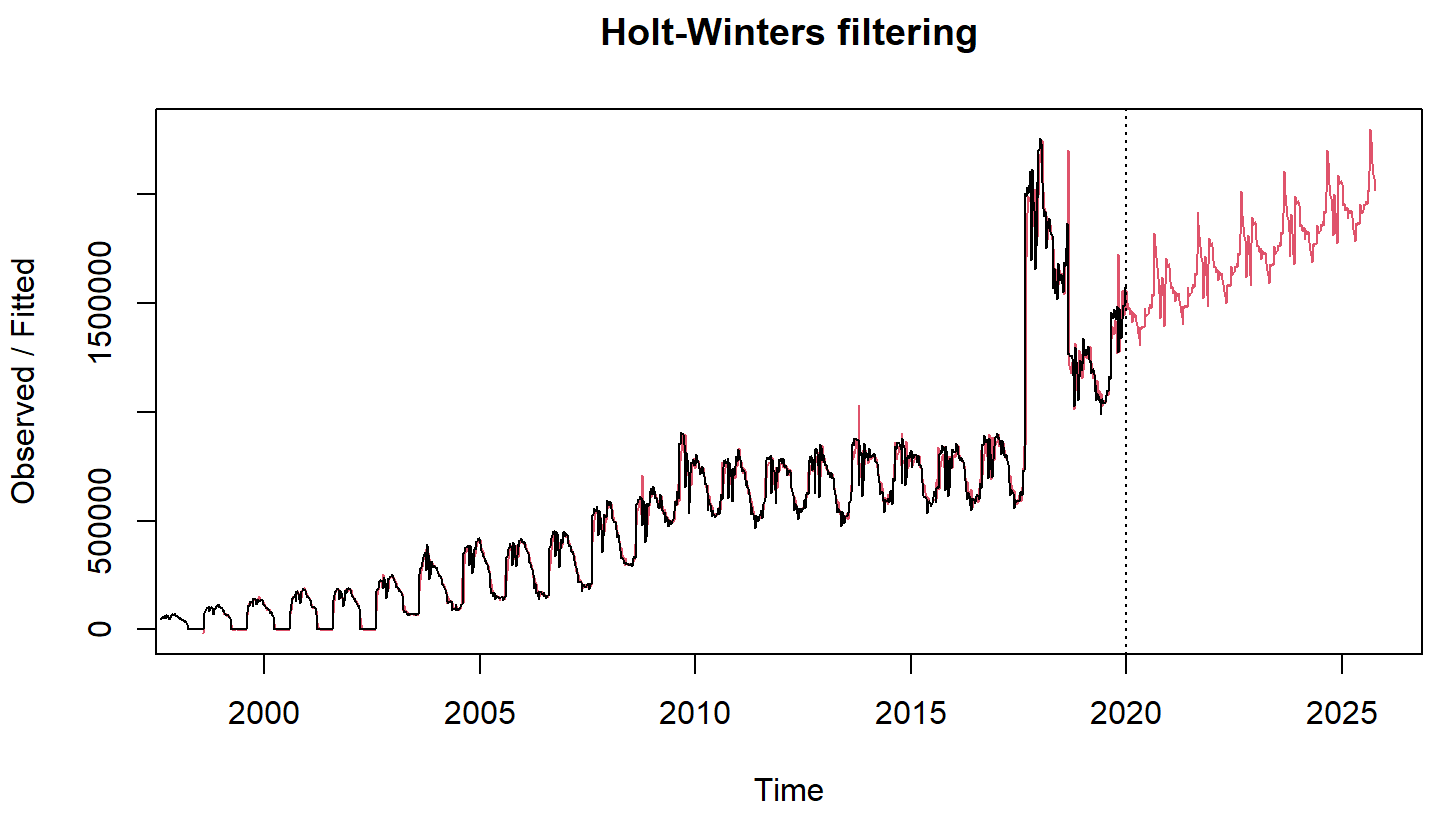
* יש לנו 0 בנתונים -השתמשנו ב "additive" מודל אדפטיבי, ולא ב "multiplicative" שהוא מכפלתי והעונתיות משפיעה עליו.



מדד אי דיוק לבחינת המודל שנבחר הפעם הוא: ערך הSSE (סכום הטעויות בריבוע של המודל):

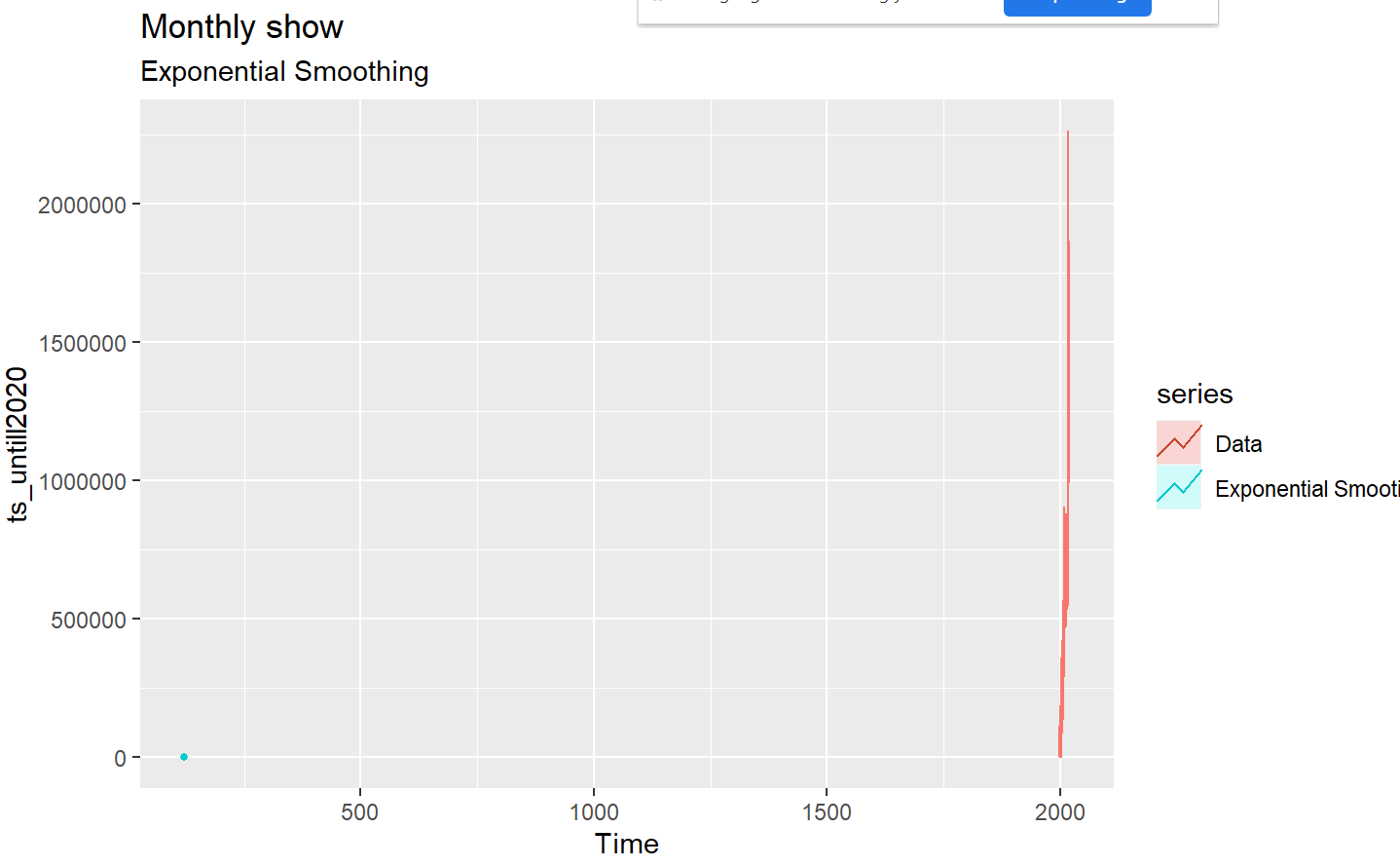
5.462384e+12

כעת נבחן את השינוי בתקופות האחרונות לאחר העליה החדשה בתוני השפעת ,ונתאים את המקדמים בהתאם ,למקדמים נקנה משקלים שונים הפעם ונסיק מהגרף:



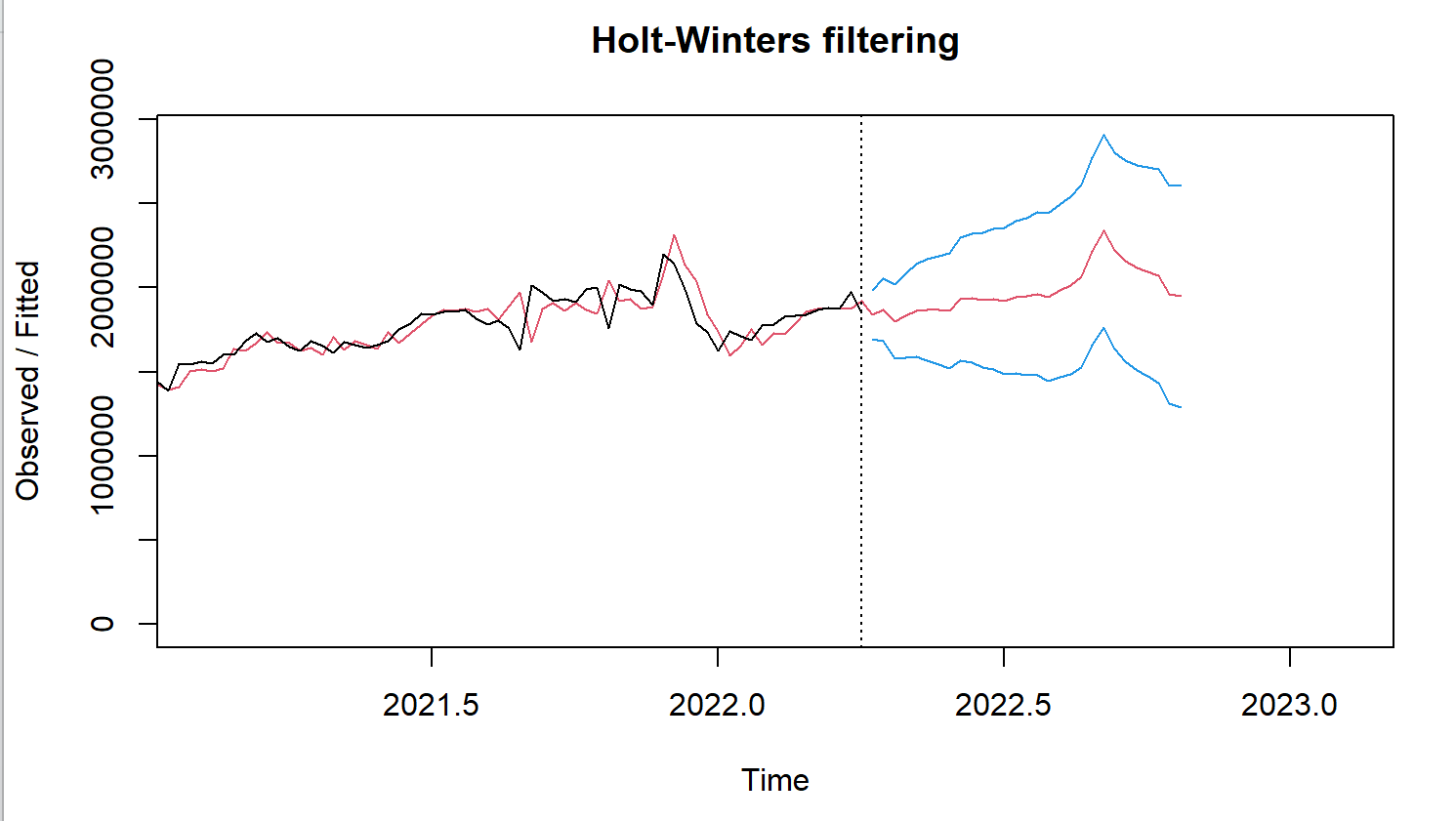
* 5.514887e+12 נקבל אי -דיוק חדש:

ננסה לחזות בעזרת החלקה מעריכית:



* **החלקה מעריכית וממוצע נע מתאימים לסדרות בלי מגמה ומכיוון שממוצע נע מקנה משקל שונה לתקופות שונות מהעבר מבחינת הנתונים ולכן לא ראינו לנכון להיעזר בו בעת שינויים במגמה ובעונתיות.**

**נבצע את מודל הולט וינטרס עם המקדמים שמצאנו מקודם על חודשים 5-10 בשנת 2022:**



תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

A=0.8

B= 0.003

G = 1

אלו הנתונים המתקבלים כעת מהמודל שבחרנו במיטבי ומתאים לחיזוי המתבסס עונתיות מבחינת השגיאה/האי-דיוק בהתייחס לדאטה: