TO DO LIST:

יהיו 2 טבלאות.

* ליצור מבנה נכון לכל טבלה, כולל הגדרת ערכים, מפתחות, לוודא שמחיקה באחד משפיעה על מחיקה באחר וכו'.
* ליצור db שמכיל את 2 הטבלאות (הריקות, כרגע) הנ"ל.

לאחר שנוצר הDB:

נרצה להתחיל להכניס אליו מידע. יש להכניס מידע ל2 הטבלאות: metadata וcontig.

עבור טבלת metadata:

* יש לגשת לקבצים שנמצאים במובקס, ולהכניס אותם בצורה נכונה למסד הנתונים שלנו. לוודא שהעמודות זהות – עלול להיות שוני!
* אם יש ערכים ריקים, להכניס none.
* לדעת לקשר בצורה נכונה לטבלת קונטיג.

עבור טבלת contig:

* על מנת לאחד את הטבלאות לטבלה אחת, נוסיף כל טבלה למסד הנתונים שלנו. כאשר מוסיפים, יש לשים לב:
  + עלינו להוסיף 2 עמודות בהתחלה: series number, sample number.
  + יש לפעמים שונים בין סוגי העמודות, צריך לראות איך מטפלים בכך, ומוודאים ערכי none בעמודות הריקות.
  + לוודא הכנסה נכונה של ערכים לעמודה הנכונה.

המלצה:

להתחיל בקטן: לבחור ניסוי אחד (GSE114724 למשל) וליצור ממבנה דאטה בייס נסיוני, עליו נעשה את כל הבדיקות: לוודא הכנסה נכונה, ערכי none, מחיקות שעובדות כמו שצריך וכו'.

לבסוף:

לכתוב (סקריפ?) קוד שמטפל בהכנסות/מחיקות חדשות לדאטהבייס באופן עצמאי. מומלץ לשלב לאחר שביצענו את "ההתחלה בקטן".

עבור ניסוי GSE176021 יש לעבוד בצורה ידנית.