

سنبدأ بنظرة عامة مفاهيمية على مستند PDF بسيط. تم تصميم هذا الفصل ليكون توجيهًا مختصرًا قبل الغوص في مستند حقيقي وإنشاءه من البداية.

يمكن تقسيم ملف PDF إلى أربعة أجزاء: الرأس والجسم والجدول الإسناد الترافقي والمقطورة. يضع الرأس الملف كملف PDF ، حيث يحدد النص المستند المرئي ، ويسرد جدول الإسناد الترافقي موقع كل شيء في الملف ، ويوفر المقطع الدعائي تعليمات حول كيفية بدء قراءة الملف.

رأس الصفحة هو ببساطة رقم إصدار PDF وتسلسل عشوائي للبيانات الثنائية. البيانات الثنائية تمنع التطبيقات الساذجة من معالجة ملف PDF كملف نصي. سيؤدي ذلك إلى ملف تالف ، لأن ملف PDF يتكون عادةً من نص عادي وبيانات ثنائية (على سبيل المثال ، يمكن تضمين ملف خط ثنائي بشكل مباشر في ملف PDF).

لأختر הכותרת והגוף מגיע טבלת הפניה המקושרת. הוא מתעדת את מיקום הבית של כל אובייקט בגוף הקובץ. זה מאפשר גישה אקראית של המסמך, ולכן בעת עיבוד דף, רק את האובייקטים הנדרשים עבור דף זה נקראים מתוך הקובץ. זה עושה מסמכי PDF הרבה יותר מהר מאשר קודמיו PostScript, אשר היה צריך לקרוא את כל הקובץ לפני עיבוד זה.