بعض الاستنتاجات من dataset التي تم العمل عليها:

* عدد الاعمدة الكلي هو 13 وعدد الاسطر الكلي هو 148654.
* الاعمدة التي تحوي على قيم فارغة هي BasePay, OvertimePay, OtherPay, Benefits,Notes and Status.
* في عملية تنظيف الداتا تم استخدام median الخاص بكل عامود لوضع هذه القيمة في القيم الفارغة او المفقودة ولم استخدم mean و ذلك لوجود فارق كبير بين كلا من ال mean و median وايضا قمت بعملية حذف لكلا من العامودين   
  Status and Notes وذلك لانهم يحتوا قيمة Nan وبالتالي لا يمكن الاستفاد منهم وحتى االعامود الخاص ب Agencyيحوي فقط على قيمة واحد وهي San Francesco وبالتالي يمكن حذفها ايضا.
* في عملية عرض ال Pie Chart الخاص بالاقسام لقد قمت بالعيديد من عمليات التوحيد حتى اتمكن من رسم Pie Chart كثر دقة ولكن يوجد الكثير من العنوانين التي لم اتمكن من ان احدد لاي قسم تنمتمي وفي الغالب اسخراج الفيتشر هذه بدقة اكبر واوضح تتطلب الكثير من العمل و التنسيق الاكبر مع مصدر الداتا حتى تكون الفيتشر هذه واقيعة وحيقية اكثر ولقد قمت بعد عمليات التوحيد البسيطة التي قمت بها باخذ فقط الاقسام العشرة التي تحوي اكبر عدد موظفين بناءا على عملية التوحيد التي قمت بها.
* في عملبات تجميع الداتا بناءا على قيمة فيتشر ما لقد اعتمدت على Year ومن ثم وجدت ان متوسط الرواتب في عام 2013 هو الكبروقيمته 77611.44 بيمنا في 2011 هو الاصغر قيمته 71744.10
* في تحديد الارتباط بين ال TotalPay وباقي الفيتشز وجدت علاقة ارتابط بينها وبين كلا من BasePay and TotalPayBenefits