یکی از دغدغه هاي اساسی سیاستگذاران در ایران نرخ فرزندآوري در خانوارهاي ایرانی است. هدف از این سوال تمرکز بر این موضوع است. برای پاسخ به این سوال از داده سرشماری در سال 1390 استفاده کنید (این داده از درگاه ملی ایران به دست آمده است و جدول مربوطه به همراه سوال آمده است). برای این منظور ابتدا باید مادر هر خانواده را یافت، سپس تعداد فرزاندان هر مادر را از طریق جمع کردن تعداد فرزندان پسر و دختر که زنده هستند، محاسبه کرد.

(اشکال مربوط به هر قسمت، شامل scatter plot های مربوطه، را نیز رسم کنید)

الف) جدول زیر را کامل کنید

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| تعداد فرزندان مادر | تعداد مادر | درص ازکل مادران | درصد با سواد | سن مادر | | نسبت تعداد دختران به پسران |
| متوسط | میانه |
| 1 |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |
| 8+ |  |  |  |  |  |  |
| کل |  | 100% |  |  |  |  |

ب) به کمک رگرسیون اثر سواد بر تعداد فرزندان را بررسی کرده و ضمن توصیف معنای پارامترهای تخمین زده شده، معناداری آماری پارامترهای مورد استفاده در مدلسازی خود را بررسی کنید.

ج) خطای مدل را به دست آورید. آیا می‌توان متغییر جدیدی را به مدل رگرسیون اضافه کرد که توصیف بهتری برای تعداد فرزندان به دست آورد (متغییر یا متغییرهای جدید را وارد مدل کنید و معناداری آماری آن‌ها را بحث کنید و با توجه به خطای مربوطه، برازش مدل‌های مختلف را مقایسه کنید)

د) سوال قبل را با به کارگیری ازGLM تکرار کنید.

ه) اثر متغیر سن مادر و شاغل بودن مادر را بر روی تعداد فرزندان به صورت جدا گانه بررسی کرده و همبستگی این دو متغییر را با باسوادی بررسی کنید. مقدار این همبستگی چگونه می‌تواند، پارامترهای محاسبه شده در دو سوال را تحت تاثیر قرار می‌دهد.

ح) به کمک آنالیز واریانس، تعامل نقش «استفاده از اینترنت» و «وضعیت آمادگی برای کار» را در تعداد فرزندان بررسی کنید و معناداری آن را بحث کنید.