

## الگوی خوشه‌بندی مشتریان بیمه عمر

(مطالعه موردی: یک شرکت بیمه‌ای)

مریم باش افشار<sup>۱</sup>

تاریخ دریافت: ۱۳۹۵/۱۲/۱۶

مسعود سعیدپناه<sup>۲</sup>

تاریخ پذیرش: ۱۳۹۶/۰۶/۲۱

فرشید تیره عیدوزهی<sup>۳</sup>

### چکیده

این پژوهش به مطالعه در اطلاعات مشتریان بیمه عمر به منظور دستیابی به الگوی خوشه‌بندی برای ارائه خدمات می‌پردازد. از جامعه شرکت‌های بیمه عمر، یک شرکت بیمه‌ای با اندازه نمونه ۱۰۰۰ نفر که در سال ۱۳۹۲ اقدام به خرید بیمه‌نامه عمر کرده‌اند، انتخاب شد. با استفاده از مدل‌های خوشه‌بندی داده‌کاوی، عوامل مؤثر و روابط بین آنها مورد بررسی قرار گرفت و در نهایت، نتایج حاصل از مدل‌های مختلف خوشه‌بندی با یکدیگر مقایسه شدند. با استفاده از نتایج به دست آمده، شرکت‌های بیمه‌ای می‌توانند مشتریان بیمه عمر را در دو گروه اصلی «مشتریان سودبخش» و «مشتریان دارای ریسک» طبقه‌بندی کنند و به هر یک از گروه‌ها، بسته‌های خدماتی متناسب ارائه دهند. به علاوه، متغیرهای جمعیتی همچون «جنسیت» و «سن» و متغیرهای بیمه‌ای همچون «حق بیمه سالیانه» و «ضریب فوت در اثر حادثه» به عنوان عوامل تأثیرگذار در شناسایی گروه‌های مشتریان هستند.

**واژگان کلیدی:** مشتریان بیمه عمر، داده‌کاوی، خوشه‌بندی، بیمه عمر.

۱. دانشجوی دکتری مدیریت فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد بین‌المللی قشم (نویسنده مسئول)، /

Maryam\_afshar22@yahoo.com

سازمان تأمین اجتماعی

۲. دانشجوی دکتری مدیریت فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد تهران مرکزی،

Masoud.saeidpanah@yahoo.com

۳. دانشجوی دکتری مدیریت فناوری اطلاعات، دانشگاه آزاد اسلامی، واحد بین‌المللی قشم،

Farshid.eidouzehi@gmail.com

## ۱. مقدمه

صنعت بیمه همانند سایر بخشهای اقتصادی در کشور به علت رویارویی با مسائلی همچون خصوصی سازی، تغییر و تحولات در قوانین و آزادسازی تعرفه‌ها از یک سو و پیشرفت فناوری و کاهش هزینه محاسبات از سوی دیگر، همواره در حال تغییرات بنیادین است تا خود را با شرایط و نیازهای جدید محیط وفق دهد. علاوه بر این جهانی شدن باعث سرعت بخشیدن به تغییرات بازار بیمه می‌شود. از این رو شرکتهای بیمه در رقابت سخت جهانی هستند و به طور جدی به دنبال یافتن بازارهای جدید، جذب بیمه‌گذاران کم‌ریسک و ارائه خدمات جدید با تقسیم‌بندی هدفمند مشتریان هستند.

هدف از تحقیق حاضر، بررسی و تحلیل داده‌های مشتریان بیمه عمر و شناسایی عوامل تأثیرگذار در توسعه خدمات و در پی آن افزایش سوددهی و توان مالی شرکتهای بیمه‌ای و کاهش ریسک مالی است که می‌تواند مورد استفاده تمامی شرکتهای بیمه‌ای (بیمه عمر) قرار گیرد. شرکتهای بیمه به منظور افزایش سود و کاهش ریسک می‌توانند با شناسایی سودمندترین بخش از مشتریان و پریسک‌ترین گروه مشتریان، برنامه‌های بازاریابی و توسعه خدماتی متناسب با هر گروه را ارائه کنند؛ با توجه به اینکه تاکنون تحقیقی در این خصوص صورت پذیرفته است برای شرکتهای بیمه‌ای امکان تحلیل عمیقی از خدمات جدید و بالقوه روی بخش خاصی از مشتریان بیمه عمر فراهم نشده است.

## ۲. مبانی نظری

تغییرات سریع در دنیای امروز، سازمانها را با چالشهای مختلفی روبه‌رو کرده است، اما در این میان سازمانهایی موفق هستند که به کمک ابزارهای مدیریتی و فناوریهای نوین، از فرصتهای ایجادشده به نفع خود استفاده کنند. در عصری که بیشتر بازارها بلوغ یافته‌اند و مشتریان جدید به سختی به دست می‌آیند، شرکتهای نیاز مبرمی به حفظ و نگهداری مشتریان و وفادار کردن آنها دارند. همراه با ارتقا و پیچیدگی بیشتر زندگی

انسان امروز، بیمه‌ها نیز از زوایای مختلف و با شتاب توسعه یافته‌اند. بیمه‌های زندگی نیز پیشرفته‌های شگرفی یافته و مجموعه متنوعی از انواع این نوع بیمه را ارائه کرده است.

با توجه به انواع ریسکها، قراردادهای بیمه نیز انواع متعددی دارند. برای تقسیم‌بندی بیمه‌ها روشهای مختلفی ارائه شده است. یک طبقه‌بندی بر مبنای کاربرد بیمه در زمینه‌های مختلف به صورت

۱. بیمه‌های اجتماعی (اجباری)، و

۲. بیمه‌های بازرگانی (اختیاری)

انجام گرفته و در عین حال مبین وجود تفاوت‌های حقوقی بین آنها نیز هست. انواع بیمه‌های بازرگانی (اختیاری) به تفکیک شامل موارد زیر است:

**الف. بیمه‌های زندگی (عمر):** بیمه عمر یکی از انواع بیمه اشخاص است. در این طرح افراد «بیمه‌گذار» با پرداخت مبلغی تحت عنوان حق بیمه، علاوه بر تأمین منابع لازم برای تشکیل سرمایه‌ای در آینده برای خود، از امکانات پوششهای بیمه‌ای متنوعی از قبیل حادثه و فوت، نقص عضو، ازکارافتادگی و یا بیماریهای خاص نیز برخوردار خواهد شد. رایج‌ترین بیمه‌نامه انفرادی در این رشته «بیمه‌نامه عمر و پس‌انداز» است که علاوه بر جنبه پس‌انداز و تشکیل سرمایه، خطر فوت را تحت پوشش قرار می‌دهد.

**ب. بیمه‌های غیرزندگی:** سایر رشته‌های بیمه به‌جز بیمه‌های زندگی، تحت عنوان «بیمه‌های غیرزندگی» دسته‌بندی می‌شوند. انواع اصلی رشته‌های بیمه غیرزندگی عبارت‌اند از: بیمه آتش‌سوزی، بیمه باربری، بیمه حوادث راننده، بیمه بدنه اتومبیل، بیمه شخص‌ثالث و مازاد، بیمه درمان، بیمه کشتی، بیمه هواپیما، بیمه مهندسی، بیمه پول، بیمه‌های مسئولیت مدنی، بیمه اعتبار، بیمه نفت و انرژی و سایر انواع بیمه‌ها.

**داده‌کاوی:** دانش داده‌کاوی فرایند کشف دانش پنهان درون داده‌هاست که با برخورداری از دامنه وسیع زیرزمینه‌های تخصصی با توصیف، تشریح، پیش‌بینی و کنترل پدیده‌های گوناگون، امروزه دارای کاربرد بسیار وسیع در حوزه‌های مختلف

است. این ابزار سایر سیستمهای پشتیبانی تصمیم‌گیری را کامل کرده و گاهی جایگزین آنها می‌شود. دو هدف اصلی داده‌کاوی، پیشگویی و توصیف است (تقوی فرد و همکاران، ۱۳۸۶).

**خوشه‌بندی:** خوشه‌بندی روشی از داده‌کاوی است که خوشه‌های معنی‌دار یا پُر استفاده از آیت‌هایی که خصوصیات مشابه دارند را می‌سازد و از روشهای خودکار بهره می‌جوید که با روش طبقه‌بندی متفاوت است. داده‌ها خوشه‌بندی می‌شوند و پس از آن با استفاده از الگوریتمهای درخت تصمیم که یکی از مهم‌ترین شیوه‌های داده‌کاوی برای کشف رابطه بین متغیرها در سیستمهای مدیریتی است به کشف رفتار خوشه‌ها پرداخته می‌شود (شهرابی و شکورنیا، ۱۳۸۶).

### ۳. پیشینه تحقیق

کجوسوسکی<sup>۱</sup> (۲۰۱۲)، به بررسی عوامل مؤثر بر تقاضای بیمه عمر در ۱۴ کشور اروپای مرکزی و جنوب شرقی با استفاده از روش اثرات ثابت طی دوره ۱۹۹۸ تا ۲۰۱۰ پرداخت. نتایج این تحقیق نشان داد که تولید ناخالص داخلی سرانه، تورم، مخارج سلامت، سطح تحصیلات و نقش قانون قوی‌ترین پیش‌بینی برای استفاده از بیمه عمر است. نرخ بهره واقعی، کنترل فساد و اثربخشی دولت به نظر نمی‌رسد که بر تقاضای بیمه عمر در کشورهای اروپای مرکزی و جنوب شرقی خیلی اثرگذار باشد.

فاین<sup>۲</sup> و همکاران (۲۰۱۱) در مطالعه خود به بررسی عوامل تعیین‌کننده حق‌بیمه عمر و غیرعمر در ۹۰ کشور درحال توسعه و توسعه‌یافته طی دوره ۲۰۰۰ تا ۲۰۰۸ پرداختند. نتایج نشان داد که درآمد سرانه، تعداد جمعیت و ساختار جمعیتی، توزیع درآمد، سیستم بازنشستگی، مالکیت دولتی شرکتهای بیمه، در دسترس بودن اعتبار خصوصی و مذهب بر بیمه عمر تأثیر دارد و بخش غیرعمر به وسیله این متغیرها و

---

1. Kjosevski

2. Feyen

متغیرهای دیگر تحت تأثیر قرار می‌گیرد. همچنین نتایج نشان می‌دهد که تعدادی از متغیرهای سیاستی نیز بر توسعه بیمه عمر تأثیر دارد.

چن<sup>۱</sup> و همکاران (۲۰۰۷)، به تحلیل عوامل مؤثر بر تقاضای بیمه عمر در ۳۰ کشور عضو سازمان همکاری اقتصادی و توسعه (OECD)<sup>۲</sup> طی دوره ۱۹۹۳ تا ۲۰۰۰ پرداختند. نتایج آنها نشان داد که درآمد، سطح تحصیلات و درجه رقابت رابطه مثبتی با تقاضای بیمه عمر دارد. مخارج تأمین اجتماعی، تورم، نرخ بهره واقعی تقاضا برای بیمه عمر را در کشورهای OECD کاهش می‌دهد.

هوانگ و گرین فورد<sup>۳</sup> (۲۰۰۵)، به بررسی تعیین‌کنندگان تقاضا برای بیمه عمر در کشورهای چین، هنگ‌کنگ و تایوان پرداختند. نتایج این مطالعه نشان داد که درآمد، تحصیلات و توسعه اقتصادی عوامل مؤثر بر تقاضای بیمه عمر هستند. ساختار اجتماعی و سیاستها تأثیر منفی بر تقاضای بیمه عمر دارند و امنیت اجتماعی و قیمت اثر معنی‌داری بر تقاضای بیمه عمر ندارند.

بصری و همکاران (۱۳۹۰)، در پژوهشی بر اساس سرمایه انسانی، عوامل اجتماعی و عوامل روان‌شناختی، الگویی را برای تقاضای بیمه عمر ارائه کردند. یافته‌های پژوهش بر اساس الگوی دومرحله‌ای نشان داد، احتمال خرید بیمه عمر در میان زنان بیشتر از مردان است. همچنین، وضعیت سلامتی افراد با تمایل به خرید بیمه عمر ارتباط منفی و معنی‌داری دارد. اما متغیرهای سرپرستی و شناخت بیمه عمر در تمایل به خرید بیمه عمر افراد تأثیر مثبتی داشته است. افزایش سن و وجود فرزند کمتر از ۱۸ سال موجب خرید بیمه عمر کمتری توسط افراد شده است. متغیرهای درآمد، امید به زندگی و انگیزه ترک ارت با سطح بیمه عمر خریداری‌شده رابطه مثبتی داشته است. سجادی و غلامی (۱۳۸۶)، با استفاده از روش کمترین توانهای دوم معمولی به بررسی رابطه میان متغیرهای کلان اقتصادی و تقاضا برای بیمه عمر در ایران پرداختند.

1. Chen

2. Organisation for Economic Co-operation and Development

3. Hwang and Greenford

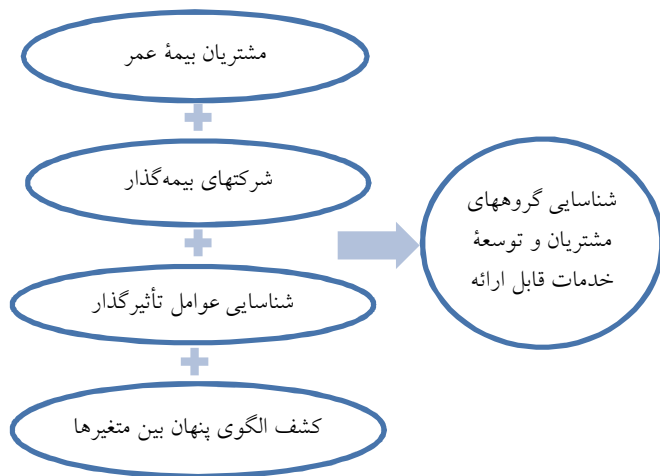
اساسی‌ترین یافته‌های این مطالعه نشان می‌دهد که تولید ناخالص داخلی و همچنین تورم پیش‌بینی‌شده، رابطه مؤثری با تقاضای بیمه عمر دارند. نتیجه جالب، رابطه مثبت تقاضا با نرخ بیمه است که متناقض به نظر می‌رسد. ضمن آنکه متغیر امید به زندگی کاملاً مؤثر بر تقاضا و با ضریب مثبت ارزیابی شده است.

عزیزی (۱۳۸۵) به بررسی رابطه میان متغیرهای کلان اقتصادی و تقاضا برای بیمه عمر در ایران طی دوره زمانی ۱۳۶۹-۱۳۸۳ پرداخت. یافته‌های تحقیق نشان می‌دهد که نرخ سپرده‌های بلندمدت، قیمت بیمه، درآمد، بازده سهام بورس اوراق بهادار و امید به زندگی از جمله مهم‌ترین عوامل مؤثر بر تقاضا برای بیمه عمر در ایران است. در این راستا قیمت بیمه رابطه معکوس قوی و مهمی با تقاضا برای بیمه عمر دارد.

رجیبیان و مهرآرا (۱۳۸۴)، تابع تقاضای بیمه عمر را با استفاده از داده‌های سری زمانی ایران طی دوره ۱۳۴۵-۱۳۸۲، مبتنی بر رویکرد  $ARDL^1$  و همچنین داده‌های پانلی برای کشورهای صادرکننده نفت با درآمد متوسط برای دوره ۱۹۹۸-۲۰۰۰ برآورد کردند. بر طبق نتایج این مطالعه، عمده‌ترین عوامل تعیین‌کننده تقاضای بیمه عمر در ایران، درآمد، میزان تحصیلات و بار تکفل‌اند. همچنین درآمد، احتمال مرگ سرپرست خانواده و تحصیلات، از مهم‌ترین عوامل مؤثر بر تقاضای بیمه‌های عمر در کشورهای نفت‌خیز با درآمد متوسط هستند. کششهای درآمدی تابع تقاضای بیمه عمر نشان می‌دهند که این محصول، در کشورهای نفت‌خیز با درآمد متوسط، لوکس و در ایران ضروری است. در واقع، افزایش درآمد، نقش اساسی‌تری در گسترش و تحولات بیمه‌های عمر در سایر کشورهای صادرکننده نفت نسبت به ایران داشته است.

#### ۴. مدل مفهومی تحقیق

بر اساس آنچه در پیشینه تحقیق بیان شد، چارچوب مفهومی تحقیق مطابق شکل ۱ طراحی شد.



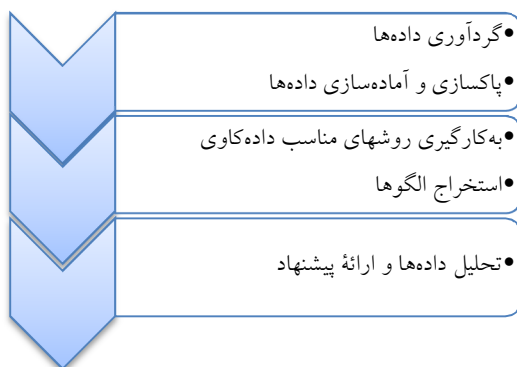
شکل ۱. مدل مفهومی تحقیق

هدف از تحقیق حاضر، بررسی و تحلیل داده‌های مشتریان بیمه عمر و شناسایی الگویی برای خوشه‌بندی مشتریان به منظور توسعه خدمات جدید و در پی آن افزایش سوددهی و توان مالی شرکتهای بیمه‌ای است که می‌تواند مورد استفاده تمامی شرکتهای بیمه‌ای (بیمه عمر) قرار گیرد. شرکتهای بیمه می‌توانند قابلیت سوددهی خود را با شناسایی سودمندترین بخش از مشتریان و اولویت‌دهی به برنامه‌های بازاریابی و توسعه خدماتی بر همین اساس افزایش دهند. با توجه به هدفهای تحقیق و الگوی عمومی پیشنهادی، سؤالاتی به شرح ذیل تعریف شد:

۱. مشتریان بیمه عمر در چند خوشه اصلی قرار می‌گیرند؟
۲. متغیرهای تأثیرگذار بر توسعه خدمات بیمه عمر کدام‌اند؟

## ۵. روش شناسی تحقیق

از آنجایی که هدف، کشف رابطه بین «مشتریان بیمه عمر» و «عوامل تأثیرگذار در ارائه خدمات» است، تحقیق از نظر هدف «کاربردی» و از نظر گردآوری اطلاعات از نوع «همبستگی» است. در تحقیق حاضر با استفاده از داده‌کاوی در اطلاعات مشتریان بیمه عمر به روش خوشه‌بندی، مطابق شکل ۲ به جستجوی عوامل تأثیرگذار در ارائه خدمات پرداخته شده است.



شکل ۲. روش تحقیق

جامعه آماری تحقیق «همه مشتریانانی که در سال ۱۳۹۲ اقدام به خرید بیمه عمر یکی از شرکتهای بیمه کرده‌اند»، در نظر گرفته شد و از آنجا که گردآوری اطلاعات از همه افراد جامعه آماری عملاً غیرممکن بود حتی اگر امکان‌پذیر باشد، به لحاظ زمان، هزینه و سایر امکانات مورد نیاز مقدور نیست، نمونه ۱۰۰۰ نفری به‌عنوان نمونه آماری تحقیق مورد مطالعه قرار داده شد. نمونه‌گیری به روش تصادفی ساده انجام شد تا به این ترتیب بتوان نتایج مشتریان نمونه انتخاب‌شده را به مشتریان کل جامعه آماری تعمیم داد.

روش انجام تحقیق از طریق مطالعه و بررسی کتب، مقالات، پایان‌نامه‌های انجام‌شده، پروژه‌های تحقیقاتی صورت‌گرفته و اینترنت است. برای تعیین متغیرهای مورد استفاده در تحقیق، فرمهای پیشنهاد بیمه و نظرسنجی از خبرگان بیمه‌ای مورد بررسی قرار گرفتند.



**بررسی فرمهای پیشنهاد بیمه:** پس از مطالعه فرمهای پیشنهاد بیمه و جوابهای داده‌شده توسط بیمه‌گذار، مشخص شد پاسخ به برخی سؤالات الزامی و برخی دیگر اختیاری است. در نگاه کلی فرمها شامل پرسشهایی در مورد مشخصات بیمه‌گذار، بیمه‌شده، استفاده‌کنندگان از سرمایه بیمه، سؤالات عمومی در مورد وضعیت سلامتی بیمه‌شده و والدین بیمه‌شده و درنهایت سوابق بیمه‌ای است. از میان پاسخهای داده‌شده برخی موارد ضروری‌تر بوده و در عقد قرارداد بیمه تأثیرگذارتر هستند.

**نظرسنجی خبرگان بیمه:** با توجه به اینکه تنها مطالعه فرمهای پیشنهاد بیمه و پاسخها نمی‌تواند ملاک مناسبی برای تعیین متغیرهای مورد استفاده در تحقیق باشد، لذا نظر ۲۰ نفر از خبرگان بیمه‌ای نیز بررسی شد. بدین ترتیب که فرمهایی شامل متغیرهای پیشنهادی با توجه به فرمهای پیشنهاد بیمه، به خبرگان تحویل داده شد و از آنها خواسته شد نسبت به امتیازدهی با توجه به میزان تأثیرگذاری متغیرها در قراردادهای بیمه عمر، اقدام کنند. درنهایت، تمامی پاسخ خبرگان مورد بررسی و تحلیل قرار گرفت و متغیرها با میانگین بیش از ۵۰ درصد، به‌عنوان متغیرهای نهایی مورد استفاده در مدل، تعیین شدند. نتایج به‌دست‌آمده، به شرح جدول ۱ است.

جدول ۱. نظرسنجی خبرگان بیمه (بر حسب درصد)

متغیرها	فوت در حادثه	سلامتی والدین	وزن و قد	فعالتهای جانبی	استان محل سکونت	سوابق بیمه‌ای	مصرف سیگار	مشروبات الکلی	مواد مخدر
میانگین نظرها	۷۲	۵۸	۵۴	۵۳	۵۰	۴۹	۴۶	۳۱	۳۱
متغیرها	سن	حق بیمه سالانه	جنسیت	درآمد	سرمایه فوت	شغل	مدت بیمه	سلامتی بیمه‌شده	تأهل
میانگین نظرها	۹۴	۹۱	۸۹	۸۷	۸۶	۸۰	۷۹	۷۶	۷۵

## ۶. تحلیل داده‌ها و یافته‌های تحقیق

برای تحلیل داده‌های خام مشتریان بیمه عمر، تک‌تک متغیرهای موجود بررسی، مرتب و پاکسازی شد. با استفاده از نرم‌افزار داده‌کاوی کلمنتاین، مدل‌های مختلف خوشه‌بندی از جمله k-means و C5 بر روی داده‌ها اعمال شد تا با مقایسه نتایج حاصل از هر مدل بتوان به بررسی روابط درونی متغیرها و میزان تأثیرگذاری هر متغیر و درنهایت به ارائه الگوی نهایی، دست یافت.

### ۶-۱. یافته‌های توصیفی تحقیق

مطابق جدول ۲، داده‌های دریافتی را به طور کلی می‌توان در سه بخش به صورت زیر قرار داد:

۱. اطلاعات جمعیتی بیمه‌شدگان شامل سن، جنسیت، وضعیت تأهل، شغل، استان و درآمد؛
۲. اطلاعات بیمه‌نامه شامل حق بیمه سالانه، سرمایه فوت، مدت بیمه‌نامه و ضریب فوت در اثر حادثه؛
۳. اطلاعات تکمیلی بیمه‌شدگان شامل وضعیت توده بدنی، میزان درآمد، سابقه بیماری و جراحی، وضعیت سلامتی والدین، شغل، استان محل سکونت و درصد اضافه نرخ که به تناسب وضعیت سلامتی فرد بیمه‌شده تعیین می‌شود.

جدول ۲. یافته‌های توصیفی

متغیرهای جمعیتی							
حداقل سن	حداکثر سن	میانگین سنی	مجرد	متاهل	زن	مرد	فعالیت‌نداشتن جنبی
۲۱ سال	۶۷ سال	۳۹ سال	٪۲۶/۳۶	٪۷۳/۶۴	٪۳۵/۷	٪۶۴/۳	٪۱۸/۸۳
اطلاعات بیمه‌نامه							
حق بیمه سالانه							
٪۲۳/۲۵	مشتریان با حق بیمه سالانه کمتر از ۴۰۰۰۰۰۰ ریال						
٪۴۴/۸۹	مشتریان با $۴۰۰۰۰۰۰ < \text{حق بیمه سالانه} <= ۶۵۰۰۰۰۰$						
٪۱۲/۷۷	مشتریان با $۱۲۰۰۰۰۰۰ < \text{حق بیمه سالانه} <= ۶۵۰۰۰۰۰$						
٪۱۹/۰۹	مشتریان با حق بیمه سالانه بیش از ۱۲۰۰۰۰۰۰ ریال						

درصد	سرمایه فوت
٪۷/۹۳	مشتریان متقاضی سرمایه فوت کمتر از ۴۰۰۰۰۰۰۰ ریال
٪۹/۹۵	مشتریان با ۷۵۰۰۰۰۰۰ < سرمایه فوت <= ۴۰۰۰۰۰۰۰ ریال
٪۲۳/۲۵	مشتریان با ۱۲۵۰۰۰۰۰۰ < سرمایه فوت <= ۷۵۰۰۰۰۰۰ ریال
٪۱۵/۰۵	مشتریان با ۱۷۵۰۰۰۰۰۰ < سرمایه فوت <= ۱۲۵۰۰۰۰۰۰ ریال
٪۲۴/۰۶	مشتریان با ۲۲۵۰۰۰۰۰۰ < سرمایه فوت <= ۱۷۵۰۰۰۰۰۰ ریال
٪۴/۸۴	مشتریان با ۲۷۵۰۰۰۰۰۰ < سرمایه فوت <= ۲۲۵۰۰۰۰۰۰ ریال
٪۱۴/۹۲	مشتریان با سرمایه فوت بیش از ۲۷۵۰۰۰۰۰۰ ریال
درصد	مدت بیمه‌نامه
٪۹/۹۵	مشتریان با مدت بیمه‌نامه کمتر از ۱۲/۵ سال
٪۱۱/۸۳	مشتریان با ۱۷/۵ < مدت بیمه‌نامه <= ۱۲/۵
٪۳۳/۸۷	مشتریان با ۲۲/۵ < مدت بیمه‌نامه <= ۱۷/۵
٪۶/۵۸	مشتریان با ۲۷/۵ < مدت بیمه‌نامه <= ۲۲/۵
٪۳۷/۷۷	مشتریان با مدت بیمه‌نامه بیش از ۲۷/۵ سال
درصد	ضریب فوت در اثر حادثه
٪۴۵/۶۱	۰/۵ < ضریب فوت در اثر حادثه (این گروه تمایلی به استفاده از سرمایه فوت در اثر حادثه ندارند، ضریب حادثه فوت ۰ در نظر گرفته می‌شود)
٪۳۷/۴۵	۱/۵ < ضریب فوت در اثر حادثه <= ۰/۵ (ضریب فوت در اثر حادثه ۱ در نظر گرفته می‌شود)
٪۱۶/۹۴	۲/۵ < ضریب فوت در اثر حادثه <= ۱/۵ (ضریب حادثه فوت ۲ در نظر گرفته می‌شود)
اطلاعات تکمیلی بیمه‌شدگان	
درصد	وضعیت توده بدنی
٪۱/۶۷	مشتریان در وضعیت کد BMI=۱ به معنای لاغری یا کمبود وزن
٪۴۵/۶۱	مشتریان با BMI=۲ به معنای وضعیت نرمال توده بدنی
٪۴۷/۲۸	مشتریان با BMI=۳ به معنای اضافه وزن
٪۵/۴۴	مشتریان با BMI=۴ به معنای چاقی و وضعیت نامناسب توده بدنی
درصد	سابقه بیماری و جراحی بیمه‌شده
٪۱۵/۲	مشتریان با سابقه بیماری و جراحی
٪۸۴/۸	مشتریان بدون سابقه بیماری و جراحی
درصد	سابقه بیماری و جراحی والدین بیمه‌شده
٪۹۳/۱	والدین مشتریان با سابقه بیماری و جراحی
٪۶/۹	والدین مشتریان بدون سابقه بیماری و جراحی

درصد	میزان درآمد
۲۰/۹۷٪	مشتریان با درآمد $\leq 5000000$ درآمد
۳۹/۶۵٪	مشتریان با درآمد $5000000 < \text{درآمد} \leq 8500000$
۲۲/۸۵٪	مشتریان با درآمد $8500000 < \text{درآمد} \leq 12000000$
۱۰/۶۲٪	مشتریان با درآمد $12000000 < \text{درآمد} \leq 24000000$
۵/۹۱٪	مشتریان با بیش از $24000000$ ریال درآمد
درصد	شغل بیمه شدگان
۵۰/۱٪	آزاد
۳/۴٪	بازنشسته
۷/۷٪	خانه‌دار
۱/۳٪	دانشجو
۵/۷٪	کارگر
۳۱/۸٪	کارمند
درصد	استان محل سکونت
۲۹/۵٪	تهران (شهرستانهای استان تهران)
۱۱٪	استانهای غربی (آذربایجان غربی - شرقی - اردبیل - زنجان - کردستان - کرمانشاه - ایلام - همدان - قزوین)
۱۷/۴٪	استانهای شرقی (خراسان رضوی - شمالی - جنوبی - یزد)
۱۷/۲٪	استانهای شمالی (مازندران - گرگان - گیلان - البرز)
۱۶/۲٪	استانهای جنوبی (سیستان و بلوچستان - کرمان - کیش - بوشهر - هرمزگان - فارس - خوزستان)
۸/۷٪	استانهای مرکزی (مرکزی - قم - اصفهان - سمنان - کاشان - چهارمحال و بختیاری - کهگیلویه و بویراحمد)

## ۶-۲. خوشه‌بندی داده‌ها

به منظور تحلیل داده‌های تحقیق و دستیابی به الگوی خوشه‌بندی مشتریان بیمه عمر، مدل‌های مختلف خوشه‌بندی با ورودیهای متفاوت ایجاد شده و الگوهای به‌دست‌آمده با یکدیگر مقایسه شدند.

**خوشه‌بندی بر اساس متغیرهای جمعیتی:** در این مرحله همه متغیرهای جمعیتی در نظر گرفته شد. خوشه‌های به‌دست‌آمده شامل اطلاعات توصیفی مشتریان بوده و در کل اهداف تحقیق را به‌تنهایی پوشش نمی‌دهد و فقط خوشه‌های با متغیرهای

تأثیرگذار جمعیتی ارائه می‌دهد. سه متغیر وضعیت تأهل، جنسیت و سن تأثیرگذارترین متغیرهای جمعیتی شناخته شدند.

در مرحله بعد متغیرهای جمعیتی تأثیرگذار به‌عنوان ورودی برای خوشه‌بندی داده‌ها انتخاب شدند. با توجه به اندازه داده‌ها، تعداد ۳ تا ۶ خوشه را تعریف کردیم. از میان مدل‌های مختلف ساخته‌شده، مدل دومرحله‌ای با تعداد ۶ خوشه و اعتبار ۰/۷۳، بهترین حالت خوشه‌بندی را نشان داد. جزئیات خوشه‌های به‌دست‌آمده از مدل در جدول ۳ نشان داده شده است.

جدول ۳. جزئیات خوشه‌ها

شماره خوشه	شرح	درصد
خوشه ۱	مردان مجرد با میانگین سنی ۲۸	۱۸/۹
خوشه ۲	مردان متأهل با میانگین سنی ۵۰	۲۷/۴
خوشه ۳	مردان متأهل با میانگین سنی ۳۲	۲۶/۷
خوشه ۴	زنان متأهل با میانگین سنی ۵۳	۱۰/۳
خوشه ۵	زنان متأهل با میانگین سنی ۳۳	۱۰/۸
خوشه ۶	زنان مجرد با میانگین سنی ۳۴	۵/۸

**خوشه‌بندی بر اساس همه متغیرها:** خوشه‌بندی با متغیرهای اصلی بیمه‌نامه و بیمه‌شده به‌نوعی اهداف اصلی تحقیق را برآورده می‌سازد. بی‌هنجاری متغیرهای بیمه‌ای نشان داد که تأثیرگذارترین متغیرهای بیمه‌ای «حق بیمه سالیانه»، «سرمایه فوت»، «ضریب فوت در اثر حادثه» و «مدت بیمه‌نامه» هستند. از دیدگاه منطقی و نتایج مدل‌های مختلف نیز می‌توان به اهمیت تأثیرگذاری این متغیرها پی برد. در مرحله بعد با در نظر گرفتن این متغیرها به‌عنوان ورودی، خوشه‌بندی داده‌ها انجام شد و از میان تمام مدل‌های خوشه‌بندی، مدل دومرحله‌ای با اعتبار حدود ۴۵ درصد و ۶ خوشه، بهترین مدل را مشخص کرد.

تحلیل خوشه‌های به‌دست‌آمده از مدل دومرحله‌ای به شرح زیر است:

**خوشه ۱:** بیشترین پرداختی حق‌بیمه و به تناسب آن سرمایه فوت بالایی را دارند. ضریب فوت در اثر حادثه ۱ است. ۱۸/۱ درصد اندازه نمونه را شامل می‌شود. با توجه به اینکه بیشترین میزان پرداختی حق‌بیمه را دارند اما دریافتی آنها به‌نسبت، کمترین است؛ لذا جزو گروه‌های سودبخش برای شرکت بیمه است.

**خوشه ۲:** حق‌بیمه پرداختی و به تناسب آن سرمایه فوت متوسطی را شامل می‌شود. ضریب فوت در اثر حادثه ۲ است. ۱۷/۷ درصد اندازه نمونه را شامل می‌شود.

**خوشه ۳:** حق‌بیمه پرداختی و به تناسب آن سرمایه فوت متوسط به بالا را شامل می‌شود. ضریب فوت در اثر حادثه ۳ است. ۸/۴ درصد اندازه نمونه را شامل می‌شود.

**خوشه ۴:** حق‌بیمه پرداختی و به تناسب آن سرمایه فوت متوسط به پایین را شامل می‌شود. ضریب فوت در اثر حادثه ۱ است. ۲۱/۵ درصد اندازه نمونه را شامل می‌شود.

**خوشه ۵:** حداقل پرداختی حق‌بیمه و به تناسب آن حداقل سرمایه فوت را دارد. ضریب فوت در اثر حادثه ۱ است. ۱۹/۹ درصد اندازه نمونه را شامل می‌شود.

**خوشه ۶:** حداقل پرداختی حق‌بیمه و به تناسب آن حداقل سرمایه فوت را دارد. ضریب فوت در اثر حادثه ۳ است. ۱۴/۴ درصد اندازه نمونه را شامل می‌شود. با توجه به آنکه حداقل میزان پرداختی به شرکت بیمه را دارند ولی ضریب فوت در اثر حادثه ۳ برابر پرداختی است و بیشترین مقدار خسارت را دریافت می‌کنند لذا جزو گروه‌های ریسکی برای شرکت بیمه است.

به‌جز خوشه ۱ و خوشه ۶، بقیه خوشه‌ها حد متوسط و نرمال دارند و به تناسب پرداخت به شرکت بیمه، خسارت دریافت می‌کنند.

**خوشه‌بندی بر اساس خوشه‌های سودبخش و ریسکی:** برای یافتن پاسخ اینکه چه افرادی با چه ویژگی‌هایی به سمت گروه‌های سودبخش و یا گروه‌های ریسکی می‌روند، مدلی از خوشه‌بندی که در آن متغیرهای جمعیتی و تکمیلی افراد به‌عنوان ورودی مدل و خوشه‌های سودبخش و ریسکی (خوشه ۱ و خوشه ۶) به‌عنوان هدف

مدل در نظر گرفته می‌شود. جزئیات الگوی کلی خوشه‌های سودبخش و ریسکی در جدول ۴ بیان شده است.

جدول ۴: جزئیات خوشه‌های سودبخش و ریسکی

سن	جنسیت	شغل	سابقه بیماری یا جراحی	فعالیت جنبی	وضعیت توده بدنی	درصد اضافه‌نرخ	حق بیمه پرداختی	ضریب فوت در اثر حادثه	خوشه پیشنهادی
۴۱ <=	زن	.	.	ندارد	.	.	کم	۱ <	ریسکی
۴۱ >	.	آزاد	.	.	۳ >=	.	کم	۱ <	ریسکی
۴۱ >	زن	خانه‌دار	.	.	.	دارد	کم	۱ <	ریسکی
۴۱ >	.	دانشجو، کارمند	.	.	.	.	کم	۱ <	ریسکی
۴۱ <=	مرد	.	.	ندارد	.	.	زیاد	۱ >=	سودبخش
۴۱ <=	زن	.	دارد	.	.	.	زیاد	۱ >=	سودبخش
۴۱ <=	زن	.	.	دارد	.	.	زیاد	۱ >=	سودبخش
۴۱ >	.	آزاد، کارگر، پازنشسته	.	.	.	.	زیاد	۱ >=	سودبخش
۴۱ >	.	خانه دار	.	.	.	ندارد	زیاد	۱ >=	سودبخش

با توجه به اندازه کم داده‌ها از مدل درخت تصمیم استفاده می‌کنیم. درخت تصمیم یکی از ابزارهای قدرتمند و متداول برای دسته‌بندی و پیش‌بینی داده‌ها با رویکرد یادگیری نظارت‌شده است. ساختار روندنمای<sup>۱</sup> آن کمک می‌کند تا درک بهتری از نتایج پیش‌بینی و کلاس‌بندی داشته باشیم. درخت تصمیم از یک درخت برای ساخت یک مدل پیش‌بینی (برآورد) استفاده می‌کند که مشاهدات درباره یک آیتم را به نتیجه‌گیری‌هایی درباره مقدار هدف آن آیتم می‌دهد. متغیرهای سن، جنسیت، و شغل اصلی، در مدل درخت تصمیم بیشترین تأثیر را داشته‌اند. تحلیل نهایی مدل درخت تصمیم نشان می‌دهد که نتایج و الگوهای حاصل از مدل ۷۹/۰۸ درصد صحیح و ۲۰/۹۲ درصد اشتباه بوده است. با توجه به نتایج تحلیل مدل و درصد بالای الگوهای

صحیح ارائه شده، نتایج قابل اتکا هستند. برای اطمینان بیشتر در کنار مدل درخت تصمیم از مدل‌های مکمل خوشه‌بندی نظیر قوانین انجمنی نیز برای تعیین ویژگی‌های گروه‌های سودبخش و ریسکی استفاده می‌کنیم. در این مرحله محدودیت حداکثر ۵ شرط را برای مدل در نظر می‌گیریم. نتایج حاصل از این مدل، ۱۳ قانون را به تفکیک خوشه‌های سودبخش و ریسکی مطابق جدول ۵ نشان می‌دهد.

جدول ۵. الگوهای حاصل از قوانین انجمنی

نوع خوشه	جزئیات خوشه
سودبخش	فعالیت جنبی داشته باشد / توده بدنی = ۲
ریسکی	شغل اصلی کارمند / کد استان = ۱ / توده بدنی = ۳
سودبخش	شغل اصلی آزاد / توده بدنی = ۳ / دارای کارت پایان خدمت
ریسکی	شغل اصلی کارمند / کد استان = ۱ / توده بدنی = ۳ / متأهل
ریسکی	شغل اصلی کارمند / کد استان = ۱ / دارای کارت پایان خدمت / متأهل
سودبخش	شغل اصلی آزاد / توده بدنی = ۳ / دارای کارت پایان خدمت / مرد
ریسکی	شغل اصلی کارمند / کد استان = ۱ / دارای کارت پایان خدمت / مرد
ریسکی	شغل اصلی کارمند / کد استان = ۱ / توده بدنی = ۳ / مرد
ریسکی	شغل اصلی کارمند / کد استان = ۱ / توده بدنی = ۳ / مرد / متأهل
ریسکی	شغل اصلی کارمند / کد استان = ۱ / متأهل
ریسکی	شغل اصلی کارمند / کد استان = ۱ / مرد / متأهل
سودبخش	فعالیت جنبی داشته باشد / شغل اصلی آزاد
سودبخش	شغل اصلی آزاد / مرد / درصد اضافه نرخ هم داشته باشد

### ۶-۳. متغیرهای اعتبارسنجی خوشه‌ها

متغیرهای اعتبارسنجی، معیارهایی برای ارزیابی کیفیت روشهای خوشه‌بندی هستند. به این دلیل که خوشه‌بندی فرایندی غیرنظارتی است و کاربر در فرایند خوشه‌بندی دخالت نمی‌کند، بسیاری از روشهای آن به شدت تحت تأثیر فرضهای اولیه خود هستند و تعداد طبقات از پیش تعیین شده و یا مثالهایی وجود ندارد که نشان دهد نتایج به دست آمده از اعتبار برخوردارند یا خیر. هدف از اعتبارسنجی خوشه‌ها یافتن خوشه‌هایی است که بهترین تناسب را با داده‌های موردنظر داشته باشند. برای ارزیابی



کیفیت خوشه‌بندی، می‌توان از متغیرهای سنجش کیفیت استفاده کرد. به کمک این متغیرها می‌توان تعداد بهینه خوشه‌ها برای خوشه‌بندی را تعیین کرد به نحوی که هر خوشه دارای تراکم بالا و جدایی لازم از خوشه‌های دیگر را داشته باشد. به منظور اعتبارسنجی خوشه‌ها شاخص نیمرخ<sup>۱</sup> مورد استفاده قرار گرفته است که یکی از متداول‌ترین متغیرهای مورد استفاده برای اعتبارسنجی خوشه‌هاست.

## ۷. نتیجه‌گیری و پیشنهادها

تحلیل مدل تحقیق نشان داد که متغیرهای جمعیتی مانند «سن»، «جنسیت»، «شغل اصلی» و متغیرهای بیمه‌ای مانند «حق‌بیمه سالانه» و «ضریب حادثه در اثر فوت» می‌تواند در شناسایی گروه‌های مشتریان بیمه عمر و ارائه خدمات شرکتهای بیمه‌ای تأثیرگذار باشند.

در این تحقیق سعی بر آن بود تا با شناسایی عوامل تأثیرگذار و ارتباط میان آنها و استفاده از نرم‌افزار تخصصی تحلیل آماری و داده‌کاوی، الگوهای مناسبی برای خوشه‌بندی مشتریان بیمه عمر و سرمایه‌گذاری و ارائه خدمات متناسب با هر یک از خوشه‌ها ارائه شود. با توجه به تحقیق انجام‌شده و نتایج حاصل از آن، سؤالات اصلی تحقیق و پاسخ آنها به شرح ذیل بیان می‌شود:

### سؤال ۱. مشتریان بیمه عمر در چند خوشه اصلی قرار می‌گیرند؟

نتایج حاصل از تحقیق نشان می‌دهد، مشتریان بیمه عمر و سرمایه‌گذاری در دو خوشه اصلی سودبخش و ریسکی جای می‌گیرند. خوشه مشتریان سودبخش از آن جهت که حق‌بیمه پرداختی بالایی دارند و به نسبت دریافتی کمتری از بیمه دارند (ضریب فوت در اثر حادثه یا صفر است یا ۱) و خوشه مشتریان ریسکی برای بیمه از آن جهت که حق‌بیمه‌های پرداختی کمی دارند و در قبال آن «ضریب فوت در اثر

حادثه» بالایی دارند یا به عبارت دیگر با حداقل پرداختی خواستار بیشترین دریافتی از شرکت بیمه هستند، حائز اهمیت و بررسی است.

هرچه تعداد مشتریان با ویژگیهای گروه‌های سودبخش بیشتر باشد، درواقع به طور غیرمستقیم توان مالی شرکتهای بیمه‌ای افزایش می‌یابد؛ لذا برای جذب مشتریان با ویژگیهای مربوط باید برنامه‌ریزیهای دقیق و خدمات مازاد تشویقی صورت پذیرد. از طرف دیگر هرچه تعداد مشتریان با ویژگی گروه‌های ریسکی بیشتر باشد، بار مالی زیادی را برای شرکتهای بیمه‌ای در پی دارد.

### سؤال ۲. متغیرهای تأثیرگذار برای توسعه خدمات بیمه عمر کدام‌اند؟

بر اساس بررسیهای به‌عمل‌آمده از فرمهای پیشنهاد بیمه عمر و سرمایه‌گذاری شرکت بیمه و گردآوری نظرات ۲۰ نفر از خبرگان بیمه‌ای، متغیرهای تأثیرگذار برای توسعه خدمات بیمه عمر، شامل «سن»، «جنسیت»، «شغل»، «سابقه بیماری یا جراحی»، «فعالتهای جنبی»، «توده بدنی»، «درصد اضافه نرخی»، «حق بیمه پرداختی» و «ضریب فوت در اثر حادثه» هستند.

مهم‌ترین محدودیت تحقیق، محدودیت تعمیم نتایج است. با توجه به اینکه داده‌های استفاده‌شده مختص شرکت بیمه تحت مطالعه بوده است لذا تعمیم نتایج حاصل از آن به سایر شرکتهای بیمه‌ای، نیاز به تحقیق بیشتری دارد.

شرکتهای بیمه‌گذار می‌توانند با کمک نتایج حاصل از این تحقیق به شناسایی گروه‌های مشتریان بیمه عمر و عوامل تأثیرگذار پرداخته و متناسب با هر گروه خدمات مورد نیاز متقاضیان بیمه عمر را ارائه و توسعه دهند که گام مهمی در جهت افزایش تعداد مشتریان با حداقل ریسک و افزایش توان مالی است. در نگاه کلی محققان، دانشجویان، کارشناسان، طراحان و دست‌اندرکاران بیمه عمر می‌توانند از نتایج حاصل از این تحقیق استفاده کنند.

در پایان، پیشنهادهایی مبتنی بر یافته‌های تحقیق به صورت زیر ارائه می‌شوند:

- شرکتهای بیمه با توجه به اطلاعات دریافتی از افراد بررسی کنند که مشتری در کدام خوشه قرار می‌گیرد و با توجه به آن بسته‌های خدماتی متناسب را به مشتریان هر گروه ارائه دهند.
- برای تأمین رضایت و ایجاد انگیزه مشتریانی که ویژگیهای خوشه‌های سودبخش را دارا هستند، خدمات تشویقی و مزاد بیشتر به آنها ارائه شود.
- سرمایه‌گذاری بیشتری به منظور جلب مشتریانی با ویژگیهای خوشه‌های سودبخش، صورت پذیرد.
- برای مشتریانی که ویژگیهای خوشه‌های ریسکی را دارند، الزام معاینه توسط پزشک معتمد و در صورت لزوم تعیین اضافه نرخ صورت پذیرد.
- با توجه به دستاوردهای این تحقیق، محدودیتهای و تجارب حاصل شده و با توجه به اینکه داده‌کاوی در مشتریان بیمه عمر و به‌دست‌آوردن رابطه‌ها و الگوهای منطقی می‌تواند نقش تعیین‌کننده‌ای در تصمیم‌گیریهای کلان داشته باشد و به تبع آن افزایش توان مالی و توسعه خدمات شرکتهای بیمه‌ای را در پی داشته باشد، لذا تحقیقها در مورد سایر شرکتهای بیمه‌ای با داده‌های متفاوت از مشتریان و استفاده از روشهای دیگر داده‌کاوی، پیشنهاد می‌شود.

## منابع

۱. باصری، ب.، امام‌وردی، ق.ا.، دقیقی اصل، ع.ر. و یراتپور، م.، ۱۳۹۰. عوامل موثر بر خرید بیمه عمر (مورد مطالعه: شهر تهران). پژوهشنامه بیمه، سال ۲۶، شماره ۱.
۲. سجادی، س.ج. و غلامی، ا.، ۱۳۸۶. بررسی تأثیر متغیرهای کلان اقتصادی جمهوری اسلامی ایران بر تقاضای بیمه عمر. فصلنامه صنعت بیمه، سال ۲۲، شماره ۲.

۳. عزیزی، ف.، ۱۳۸۵. رابطه میان متغیرهای کلان اقتصادی و تقاضا برای بیمه عمر در ایران (۱۳۸۳-۱۳۶۹). فصلنامه مدرس علوم انسانی، دوره ۱۰، شماره ۴.
۴. رجبیان، ا. و مهرآرا، م.، ۱۳۸۴. تقاضا برای بیمه عمر در ایران و کشورهای صادرکننده نفت. مجله تحقیقات اقتصادی. شماره ۷۴.
۵. شهرابی، ج. و شکورنیا، و.، ۱۳۸۶. مروری بر مفاهیم، وظایف و فرایند داده‌کاوی سازمانی. اولین کنفرانس ملی داده‌کاوی ایران.
۶. تقوی‌فرد، م.ت.، منصوری، ط.، ناصرزاده، س.م.ر. و فراست، ع.ر.، ۱۳۸۶. داده‌کاوی و کاربرد آن در تصمیم‌گیریها. نشریه دانش مدیریت، دوره ۲، شماره ۷۹.
7. Kjosevski, J., 2012. The determinants of life insurance demand in central and southeastern Europe. *International journal of Economics and Finance*, 4(3), pp. 237-247
8. Feyen, E., Lester, R. and Rocha, R. 2011. What Drives the Development of the Insurance Sector. *Policy Research Working Paper no 5572*, World Bank, Washington.
9. Chen, P.F., Lee, C.C. and Lee, C.F. 2012. How does the development of the life insurance market affect economic growth? Some international evidence. *Journal of International Development*, 24(7), pp. 865-893.
10. Hwang, T. and Greenford, B. 2005. A cross-section analysis of the determinants of life insurance consumption in Mainland China, Hong Kong, and Taiwan. *Risk Management and Insurance Review*, 8(1), pp. 103-125